

جامعة مؤتة

كلية الدراسات العليا

الكشف عن الأداء التفاضلي لمتغير الجنس في اختبار PISA الدولي لعام 2012.

إعداد الطالبة

خالده محمود فارس البشابشه

إشراف

الدكتور صبري حسن خليل الطراونه

رسالة مقدمة إلى كلية الدراسات العليـــــا

استكمالاً لمتطلبات الحصول على درجـــــة

الماجستير في القياس والتقويم/ قسم علم النفس

جامعة مؤتة، 2016

الإهداء

إلى روح والدي الطاهرة

إلى أمي الغالية أمد ا في عمرها

إلى رفيق عمري وشريك حياتي الذي ساندني وكان لي خير معين

إلى أولادي آيه، سلاف، يوسف، ومحمد ......الذين تحملوا معي مشقة العمل،

وشاركوني تحقيق الهدف .... حفظهم ا ورعاهم

إلى إخوتي .... وأخواتي

إلى زملائي .... وزميلاتي

إلى طلابي .... وطالباتي

أهدي هذا الجهد المتواضع مع خالص محبتي وتقديري

خالده محمود فارس البشابشه

الشكر والتقدير

الحمد رب العالمين، والصلاة والسلام على أشرف المرسلين سيدنا محمد وعلى آله وصحبه أجمعين.

أقدم خالص شكري وتقديري لأستاذي المشرف على رسالتي الدكتور الفاضل صبري الطراونه، الذي لم يبخل علي بعلمٍ أو معرفة، فكان لي خير ناصح وخير معلم.

كما وأتقدم بجزيل الشكر والعرفان إلى أعضاء لجنة المناقشة، الأستاذ الدكتور فؤاد الطلافحه، الدكتور راجي الصرايره، والدكتور أحمد الثوابيه، لتكرمهم بمناقشة رسالتي وتقديم النصح والإرشاد.

ولا يفوتني أن أتقدم بخالص الشكر والتقدير لمن ذلل لي مهمتي بالحصول على بيانات الدراسة، الدكتور عماد العبابنه من المركز الوطني لتنمية الموارد البشرية.

وكل الشكر والتقدير لكل من ساعدني وذلل لي مهمة لإنجاز هذا العمل ....

وا ولي التوفيق.

خالده محمود فارس البشابشه

فهرس المحتويات

قائمة الجداول

قائمة الأشكال

قائمة الملاحق

الملخص

الكشف عن الأداء التفاضلي لمتغير الجنس في اختبار PISA الدولي لعام 2012

خالده محمود البشابشه

جامعة مؤتة، 2016

هدفت الدراسة الكشف عن الأداء التفاضلي لمتغير الجنس في الاختبار الدولي PISA 2012 في مبحثي العلوم والقراءة، وذلك باستخدام طريقتي مانتل – هانزل، ونسبة الأرجحية العظمى، وبلغ عدد فقرات اختبار القراءة (44) فقرة، وعدد فقرات اختبار العلوم (53) فقرة، تم تطبيق الاختبار على (7038) طالباً وطالبة من الصف العاشر، وكشفت الدراسة عن النتائج التالية:

وجود أداء تفاضلي لصالح الإناث على (39) فقرة من فقرات اختبار القراءة، وعلى (40) فقرة من فقرات اختبار العلوم، وفقرة واحدة لصالح الذكور في اختبار العلوم فقط وباتفاق جميع الاختبارات المستخدمة.

وكانت نسب الأداء التفاضلي على أبعاد المستوى المعرفي في اختبار القراءة (100%,79%,100%) في أبعاد (الوصول والاسترجاع ،الربط والتفسير، التأمل والتقييم) على الترتيب، أما بالنسبة لاختبار العلوم فكانت النسب (78%,100%,64%) في أبعاد (استخدام الأدلة العلمية، تحديد القضايا العلمية، شرح الظواهر علمياً) على الترتيب.

وتوصي الدراسة بالاستفادة من تجارب الدول الأخرى ذات الأداء المتميز، وزيادة تطلع وزارة التربية والتعليم في الأردن لردم الفجوة الجندرية في التعليم .

Abstract

Exploring Differential Item Functioning (DIF) For Gender Difference In PISA Test (2012)

Khaleda Mahmoud Al-Bashabsheh

Mu'tah University,2016

This study aimed to explore differential item functioning (DIF) for gender difference in PISA Test (2012) in reading and science tests ,utilizing Mental-Hansel ,LikeliHood- Ratio test methods, and the number of the items reached (53) item in science test ,and (44) item in reading test, Where the test was applied on (7038) students of both genders from toth grade,

In reference of the research questions ,the study revealed :

Existence of differential item functioning for females on (39) item in Reading Test, and (40) item in Science Test ,and only one item for males in Science Test, with agreement of all the tests used.

And the percentage of the differential item functioning (DIF) on dimensions of knowledge level in Reading Test were (%100,%79,%100) in dimensions of (Access and retrieve, Integrate and interpret, Reflect and evaluate) respectively, but for the Science Test the percentage were (%64,%100,%78) in dimensions of

(Using scientific evidence, Identifying scientific issues, Reflect and evaluate) respectively.

The study recommends making use of other countries high performance ,and increasing the aspiration of the Jordanian ministry of education to get rid of the educational gender gap.

الفصل الأول
خلفية الدراسة وأهميتها

1.1 مقدمة:

تعتمد دقة صنع القرار في أي مجال من المجالات الإنسانية على نوعية البيانات ودقتها ومدى شمولها للمتغيرات التي يتكون منها مجال البحث، وعادة ما تكون هذه البيانات نتيجة الأداء على أدوات القياس المختلفة، والتي تعد الاختبارات أكثرها استخداما، حيث أصبحت الاختبارات تستخدم بشكل متزايد في اتخاذ القرارات المتعلقة بالتوظيف، والتعليم والتشخيص، بهدف وضع الفرد المناسب في المكان المناسب .

ويعتمد الاستخدام الفعال للاختبارات بدرجة كبيرة على الخصائص السيكومترية للاختبار، حيث ينصب اهتمام مطوري الاختبارات على الصدق والثبات للتأكد من أن الاختبارات التربوية مقاييس صادقة لما وضعت لقياسه، وبدقة كبيرة، ولكن هناك خاصية إضافية لقيت اهتماماً واسعاً من قبل المختصين في القياس، ونقاد الاختبارات، منذ الستينات، هي خاصية الأداء التفاضلي للفقرات، والتي تحدث عندما يكون للمفحوصين من مجموعات عرقية مختلفة، احتمالات مختلفة للإستجابة الصحيحة، بعد مساواتهم في القدرة الأساسية التي يقيسها الاختبار (Crocker & Algina, 1986).

وتمثل الرياضيات والعلوم والقراءة في المرحلة الأساسية الدعامة والأساس المتين لبناء مجتمع علمي، لذا نجد أن معظم الأمم تعتمد على هذه المواضيع في تطورها العلمي وتقدمها التقني، وامتلاكها لأسباب ووسائل الريادة في العالم المعاصر (المساعفة، 2005).

وقد أدرك المعنيون بشؤون التربية والتعليم في منظمة التعاون والتنمية الاقتصادية Organisation for Economic Co-operation and Development (OECD)، أهمية الرياضيات والعلوم والقراءة، وذلك بتطبيق برنامج التقييم الدولي للطلبة (PISA) Programme for International Student Assessment))

وتكمن أهمية الدراسات الدولية للمشاركين في فهم أنظمتها التربوية بشكل أفضل، مما يساعد صانعي القرارات وراسمي السياسات التربوية من تحديد معايير واقعية للتحصيل أو الأداء التربوي والتي تعينهم في مراقبة وتقويم نجاحات أو إخفاقات نظمهم التربوية.

إن دراسات المقارنة الدولية التي من شأنها أن توفر إضافة مهمة للدراسات التي تجرى على الصعيد الوطني لكل دولة من الدول المشاركة، فهي توسع مدى الخبرة الضرورية لقياس وتقييم التحصيل التربوي ،وتوفر درجة أعلى من الثقة في تعميم الدراسات التي تفسر العوامل المهمة في التحصيل التربوي، وتزيد من احتمال انتشار أفكار جديدة تسهم في تحسين تصميم أو إدارة المدارس والصفوف، وتضيف أبعاداً جديدة للبحث التربوي في الدول المشاركة جميعها، كما توفر دراسة موضوعية، وتقيما غير متحيز للتجديدات التربوية التي تدخلها الدول المشاركة على أنظمتها التربوية من خلال خطط وبرامج التطوير التربوي التي تنفذها البلدان المشاركة في مثل هذه الدراسات، وتعين في الوقت نفسه على تلافي جوانب القصور، والضعف لهذه البرامج والإجراءات المستقبلية التي من شأنها أن تحسن أداء أنظمتها التربوية (OECD, 2012).

ويشارك الأردن في الدراسات الدولية لتقييم أداء الطلبة منذ عام 1991، وبدأت مشاركة الأردن الأولى في البرنامج الدولي لتقييم الطلبة الذي ينفذ كل ثلاث سنوات في دورة PISA في عام 2006 بمشاركة (57) دولة من دول العالم، ويعد المركز الوطني لتنمية الموارد البشرية المؤسسة الوطنية التي تشرف على الدراسة وتنفذها بالتنسيق مع الجهات الدولية المشرفة عليها والمؤسسات التربوية الوطنية المختلفة والتي تتمثل بوزارة التربية والتعليم والثقافة العسكرية ووكالة الغوث والمدارس الخاصة، إذ يقوم المركز بتنفيذ الاختبارات، وإصدار التقارير، ودعم الدراسات التربوية المستفيدة من نتائج هذه الدراسات.

ويأمل الأردن من خلال مشاركته في هذه الدراسة بتطوير مؤشرات نوعية لمدى نجاح النظام التربوي في إعداد وتهيئة الطلبة للعب دور بناء كمواطنين في مجتمعاتهم وتحديد مدى اكتساب الطلبة للمعارف والمهارات في مجالات القرائية

والرياضيات والعلوم لمساعدة النظام التربوي على تشخيص نقاط قوته وضعفه لتحسين مخرجات التعليم(المركز الوطني لتنمية الموارد البشرية، 2012).

ولقد سعى البرنامج الدولي لتقييم الطلبة في دورة عام 2012 قياس قدرة الطلبة على توظيف وصياغة وتفسير المشكلات والمعرفة الرياضية، واستخدام الأدلة العلمية وتحديد القضايا العلمية في مجال العلوم، وقياس قدرة الطالب على فهم واستيعاب واستخدام النصوص المكتوبة في مجال القراءة.

ولما كانت هذه الاختبارات تعد وتنفذ وتصدر نتائجها عن هيئة دولية وليست وطنية في الدول التي تطبق فيها هذه الاختبارات، فإنه يخشى من أن يكون تصنيف وترتيب الدول حسب نتائجها على هذه الاختبارات فيه شيء من عدم الإنصاف بسبب احتمالات وجود تحيز في فقرات هذه الاختبارات قد يؤدي إلى تشوه في مصداقية الاختبارات، والنتائج، والتصنيفات الصادرة عنها، من هنا جاءت فكرة هذه الدراسة وهي الكشف عن الأداء التفاضلي لمتغير الجنس في فقرات الاختبار الدولي (PISA) للعام 2012 .

2.1 مشكلة الدراسة وأسئلتها:

عندما تقوم مؤسسة معنية متخصصة في بناء اختبارٍ ما فإنها تقوم بمجموعة من المعالجات الإحصائية بغية التأكد من جودة الاختبار المصمم وغالباً ما تكون هذه المعالجات تدور في فلك دلالات الصدق والثبات للاختبار ككل، إن البحث في دلالات الصدق والثبات لا يعطي مؤشراً كافياً لعدالة الاختبار ومراعاة البيئات الثقافية المختلفة، وانطلاقاً مما سبق جاءت هذه الدراسة لتبحث في فقرات الاختبار الدولي (PISA) للعام 2012 للكشف عن الأداء التفاضلي لمتغير الجنس ، لذا سعت هذه الدراسة للإجابة عن الأسئلة الآتية:

1. هل يوجد أداء تفاضلي لمتغير الجنس على فقرات اختبار Pisa 2012 لمادة القراءة؟
2. هل يوجد أداء تفاضلي لمتغير الجنس على فقرات اختبار Pisa 2012 لمادة العلوم؟

3. ما هي أبعاد المستوى المعرفي التي تظهر أداء تفاضلياً علـى فقـرات اختبـار القراءة في اختبار Pisa 2012 ؟
4. ما هي أبعاد المستوى المعرفي التي تظهر أداء تفاضلياً علـى فقـرات اختبـار العلوم في اختبار Pisa 2012 ؟

3.1 أهمية الدراسة:

هدفت هذه الدراسة إلى الكشف عن الأداء التفاضلي لمتغير الجنس في فقـرات الاختبار الدولي (Pisa) للعام2012 في الأردن، حيث تقدم دراسة Pisa - باعتبارها دراسة دولية تضم دولا من كافة أنحاء العالم- بيانـات حـول النتـائج التعليميـة والتربوية لآلاف الطلبة، لذلك فإن المشاركة فـي اختبـار PISA لا تمكـن الـدول المشاركة فيه فقط، بل تتيح لها أيضا مقارنة مستويات المهارات بـين مجموعـات معينة ضمن الدولة الواحدة، ومن الاهتمامات الرئيسة لصناع القرار والتربويين في الأردن هو دراسة الفارق في الأداء واكتساب المهارات بين الطلبة الذكور والإنـاث بعمر15 سنة، وإعادة النظر في سياساتها ونظمها التربويـة ومناهجهـا التربويـة وتطوير كوادرها البشرية بناء على هذه النتائج التي حصلت عليها.

4.1 التعريفات المفاهيمية والإجرائية:

الأداء التفاضلي للفقرة: دلالة مشتقة إحصائياً للتعبير عن الفروق في احتمالات الاستجابة الصحيحة على الفقرة بين مجموعتين من المفحوصـين ممـن هـم فـي المستوى نفسه من القدرة.

اختبار (Pisa): دراسة دولية تهدف إلى معرفة مدى امتلاك الطلبة في عمـر (15) سنة للمهارات والمعارف الأساسية في مجالات العلوم والرياضيات والقراءة، وتشرف على هذه الدراسة منظمة التنمية والتعاون الاقتصادي(OECD) .

5.1 حدود الدراسة:

1- اقتصرت هذه الدراسة على نتائج طلبة الصف العاشر في اختبـارات العلـوم والقراءة (Pisa 2012) في الأردن.

2- اقتصرت هذه الدراسة على بعد واحد من أبعاد الأداء التفاضـلي وهـو الأداء التفاضلي تبعا لمتغير الجنس.

3- اقتصرت هذه الدراسة على استخدام طريقتين من طرق الكـشف عـن تحيـز الفقرات وهي طريقة مانتل -هانزل، ونسبة الأرجحية العظمى.

الفصل الثاني
الإطار النظري والدراسات السابقة

1.2 لإطار النظري

تتأثر درجات الاختبار التي يحصل عليها الأفراد بمصادر التبـاين المختلفـة، كأداة القياس نفسها، وإجراءات تطبيق الاختبار، وتصحيحه، والمصادر التي تتعلـق بالأفراد المفحوصين، إضافة إلى السمة أو الصفة التي يقيسها الاختبار، ولأن هـذه المصادر لا يمكن إلغائها أو حذفها، فمن الضروري أن لا تكون متحيزة لمجموعـة فرعية من المفحوصين ضد مجموعة فرعية أخرى، مثل الـذكور مقابـل الإنـاث (Crocker and Algina, 1986).

وهناك اختبارات دولية مقننه يركز مطوري هذه الاختبارات علـى إخراجهـا بالصورة الصحيحة والعادلة ما أمكن، ومن هذه الاختبارات اختبار PISA الدولي.

2.2 التعريف بالبرنامج الدولي لتقييم الطلبة وأهدافه:

يجمع البرنامج الدولي لتقييم الطلبة بين ثلاثة مجالات محددة وهـي القـراءة والرياضيات والعلوم، ومهارة حل المشكلات في تلك المجالات دون تركيـز علـى محتوى المناهج الدراسية المتعلقة بها، بل على المعرفة والمهارات الأساسية التـي يحتاجها الطلبة في حياتهم، إضافة إلى التركيز على استيعاب المفاهيم و القدرة على العمل في أي مجال تحت مختلف الظروف بهدف قياس مدى نجاح الطلبة الذين بلغ سنهم 15 سنة ممن هم على وشك استكمال تعليمهم الإلزامي والاسـتعداد لمواجهـة تحديات مجتمعاتهم اليومية، ويعتبر البرنامج الدولي لتقييم الطلبـة (PISA) جهـداً تعاونياً بين الأعضاء المشاركين من بلدان منظمة التعـاون والتنميـة الاقتـصادية (OECD)، إضافة إلى عدد آخر من الدول المشاركة في الاختبار.

وتنتهج هذه المنظمة أسلوبا عاما لتقييم المعارف والمهارات، بهدف توظيـف المعرفة في المهام، والتحديات اليومية؛ لتعكس هذه المهارات قـدرة الطلبـة علـى

مواصلة التعليم مدى الحياة بتطبيق ما تعلموه في المدرسة في مختلف مجالات حياتهم، وتقييم اختياراتهم وقراراتهم.

وكلمة (Pisa) هي الأحرف الأولى للجملة الانجليزية (Programme for International Student Assessment) وتعني برنامج التقييم الدولي للطلبة؛ وهو عبارة عن مجموعة من الدراسات تشرف عليها منظمة التعاون والتنمية الاقتصادية (OECD) وتجرى كل ثلاثة أعوام بهدف معرفة مدى امتلاك الطلبة من العمر (15) سنة للمهارات والمعارف الأساسية في الرياضيات والعلوم والقرائية التي تعينهم على المشاركة الفاعلة في المجتمع.

تعتمد الدراسة على معايير موحدة مثل تساوي أعمار الطلاب (15) عام فقط دون اعتبار لصفوفهم الدراسية، وتماثل الأسئلة التطبيقية، وتحييد عوامل التاريخ والثقافة المحلية، حيث لا يتم إجراء اختبارات في التاريخ أو اللغة أو الدين مثلا.

ويدير هذا البرنامج خمسة مراكز علمية وهي:

المجلس الأسترالي للبحث التربوي.

المركز الوطني الهولندي للقياس التربوي.

مركز الاختبارات في الولايات المتحدة.

المركز الوطني الياباني لدراسات السياسات اليابانية.

المركز الأمريكي للتحليل الإحصائي

ويتكون البرنامج من:

1- أدوات الدراسة فقرات الاختبار: وتحتوي على أسئلة في الرياضيات والعلوم والقرائية من نوع الاختيار من متعدد أو الإجابة القصيرة أو الطويلة. يعتمد الاختبار على عناقيد من الأسئلة يتم توزيعها على كراسات الاختبار بطريقة منظمة، والعنقود هو عبارة عن مجموعة من الوحدات تشتمل على أسئلة وضعت مع بعضها بعضا والمدة المقررة للامتحان ساعتان.

2- إستبانة الطالب: مجموعة من الفقرات التي توفر معلومات عن الخلفية الأسرية والأكاديمية للطلبة واتجاهاتهم وطموحاتهم وطرق تعلمهم للرياضيات والعلوم والقرائية وزمنها المقرر 30 دقيقة.

3- إستبانة المدرسة: مجموعة من الفقرات التي توفر معلومات عن البيئة المدرسية والهيئة التدريسية والطلبة والمناهج الدراسية وبرامج التدريب للعـاملين فـي المدرسة، والوقت الذي يقضيه الطلبة في التعلم داخـل المدرسـة، وعلاقـة المدرسة مع المجتمع المحلي وأسر الطلبة والزمن المقرر لهذه الإسـتبانة 30 دقيقة.

ويتم اختيار المدارس عشوائيا من قبل فريق دولي، والمـدارس التـي يـتم اختيارها تلبي المتطلبات الدولية في اختيار أفرادها، وهذه المتطلبات تضمن البيانات النوعيـة الجيـدة ويمكـن أن تجعـل المقارنـات واقعيـة بـين الـدول. وهناك متطلبات لاختيار المدرسة لطلابها وهي:

1. إبلاغ المدرسة التي يتم اختيارها من قبل المركز الوطني .
2. ضرورة تبليغ المعلمين والطلبة وأولياء الأمور في المدرسة التي يتم اختيارها
3. ضرورة ترتيب الطلبة حسب تاريخ ميلادهم.
4. ضرورة توفر غرفة مناسبة للاختبار.
5. ضرورة استكمال الإستبانة لكل مدرسة مشاركة.
6. الإجابة على إستبانه قصيرة للمعلم ومدير المدرسة تحتوي على بعض الأسئلة ويكون المطلوب من الطالب:

أ. الإجابة على الاختبار والذي يضم أسئلة في الرياضيات، العلوم، والقراءة.

ب. الإجابة على بعض الأسئلة حول الخلفية الدراسـية وطـرق التـدريس وعلـى اتجاهاتهم نحو الرياضيات وتجاربهم.

لقد سعى البرنامج الدولي لتقييم الطلبة (PISA) في دورة عام 2012 قياس قدرة الطلبة على توظيف وصياغة وتفسير المشكلات باستخدام الرياضيات .

أما بالنسبة لاختبار العلوم يعد امتلاك المعرفـة العلميـة والتكنولوجيـة فـي المجتمع مؤشراً على قدرة أفراد المجتمع على المشاركة الفاعلـة فـي المجتمعـات المتقدمة، لذا أصبح هذا النوع من المعرفة يمثل أولوية للمجتمعات التي ترغب فـي إعداد أفرادها إعداداً جيداً للحياة، وفي تحسين معيشة الأفراد(المركز الوطني لتنمية الموارد البشرية، 2012).

وترتبط المعرفة العلمية التي تقيمها دورة عام 2012 بالمستويات المعرفية التالية: تحديد القضايا والمسائل العلمية، تفسير الظواهر العلمية، وتوظيف الأدلة والبراهين .

وفي القراءة تم التركيز على تقييم فهم الطلبة للنصوص المطبوعة والرقمية على مختلف أنواعها مثل المقالات والنشرات والإعلانات والجداول، كما ويقيم مجال القراءة قدرة الطلبة على استخدام النصوص والتفكير فيها والتعامل معها وكانت أبعاد المستوى المعرفي هي: الوصول إلى المعلومات واسترجاعها، وربطها وتفسيرها، والتأمل والتقييم. **OECD (PISA 2012)**

3.2 التحيز والأداء التفاضلي:

لقد بدأ الاهتمام بالتحيز منذ عام 1911 عندما كاف بنية بمراجعة الصورة الأولى لاختباره الذي بناه لقياس القدرة العقلية، بعدما تبين أن متوسط أداء أبناء الطبقة الاقتصادية العالية أعلى بكثير من أداء أبناء الطبقة الاقتصادية المتدنية، وهذا الأمر أدى للتفكير بأن بعض الفروق تقيس تأثير التدريب الثقافي في الأسرة أو البيت أو المدرسة أكثر من قياسها للقدرات العقلية للأفراد المفحوصين(Camili and Shepard, 1994).

يرى هامبلتون وروجرز أن الفقرة تعتبر متحيزة، إذا كان الفرق في المساحة بين منحنيات خصائصها، في المجموعات المختلفة المتكافئة في القدرة ذات دلالة إحصائية، كأن يكونوا ذكوراً أو إناثاً، أو أن يكونوا من أعراق مختلفة، أي أن احتمالات الاستجابة الصحيحة على الفقرة مختلفة للأفراد في المجموعات الفرعية ممن لهم نفس مستوى القدرة (Hambleton & Rogers, 1995).

ويعرف ألجينا وكروكر التحيز في الفقرة على أنه: اختلاف الأداء على فقرة ما من قبل أفراد من مجموعتين فرعيتين يمتلكون نفس الموقع على متصل السمة (Crocker & Algina, 1986).

ويشير النبهان أن التحيز يعتبر من أخطاء القياس المنتظمة بمعنى أنه يؤثر في درجات المفحوصين بمقدار واتجاه ثابتين، وتعتبر الأخطاء المنتظمة أخطاء يـسهل اكتشافها وضبط أثرها في درجات المفحوصين (النبهان، 2004).

ويشير علام إلى ضرورة مراعاة من يقوم بإعداد الاختبارات المقننـة بتبـاين الأفراد المختبرين في خلفياتهم الثقافية، وخصائصهم المميزة عند انتقـاء مفـردات الاختبار لأن عدم مراعاة خصائصهم وقدراتهم يؤثر دون أدنى شـك فـي نتـائج الاختبار ويقلل من شأن هؤلاء الأفراد بطريقة أو بأخرى (علام، 2006).

ومن هنا لا بد من مراعاة أن يكون الاختبار الذي يـصمم لأي غـرض مـن الأغراض خالياً ما أمكن من الفقرات التي تعتبر فقرات متحيزة ضد مجموعة أو فئة من الفئات.

ويعتبر عوده التحيز بجميع أنواعه من أبرز المشكلات والآثار السلبية التي تترتب عليها نتائج الاختبارات وذكر منها:

1. تحيز عينة الأسئلة.

2. تحيز صياغة الفقرات فقد تكون الصياغة لصالح فئة دون أخرى.

3. التحيز في محتوى الفقرة (عوده، 2004).

استخدم مفهوم لأداء التفاضلي(Differential Item Function, DIF) في الأدب التربوي منذ بداية الثمانينات للتعبير عـن الطرائـ ـق والمعالجـ ـات الإحـصائية المستخدمة للكشف عن تحيز فقرات الاختبارات، فهي دالـة مـشتقة إحـصائيا للتعبـير عـن الفـرق في الاسـتجابة للفقرة بين مجموعتين عند مـستوى قـدرة واحد (Camili & Shepard, 1994).

ويعرف هامبلتون الأداء التفاضلي للفقرة سيكومتريا بأنه الفرق في احتمـالات الاستجابة الصحيحة للفقرة، في المجموعـات المختلفـة والمتـساوية فـي القـدرة (Hambleton, 1991).

أما هامبلتون و سوامينثان وروجرز فقـد عرفـوا الفقرة التـي تظهـر أداء تفاضليـاً، بأنهـا تلـك الفقرة التي تكـون دالـة الاستجابـ ـة عليهـ ـا غيـ ـر متماثلة في المجموعات الفرعية المختلفـة، أما إذا كانت دوال استجابة الفقرة متماثلة

فـي المجموعـات المختلفـة فـإن الفقـرة لا تظهـر أداء تفاضليـاً (Hambleton & Swaminathan & Rogers, 1991) .

ويفرق دورنز وهولاند بين مفهوم التحيز ومفهـوم الأداء التفاضـلي للفقـرة ،فالأداء التفاضلي للفقـرة يدرس الخصائص السيكومترية للفقـرة مـن حيـث الكيفية التي تعمـل بها الفقرة في المجموعـة بشكل مختلـف عـن المجموعـة الأخرى، بينما تحيز الفقـرة يحمـل المعنى الاجتماعي، وعدم العدالة والمـساواة (Dorans & Holland, 1993).

ويرى وليمز أن الفقرة تبدي أداء تفاضـلياً إذا كانـت احتمـالات الإسـتجابة الصحيحة عنها مختلفة للمفحوصين المتساوين في السمة المقاسة فـي المجموعـات المختلفة، بينما تعد الفقرة متحيزة عندما تقيس غير ما هو متوقع أن تقيسه أو عندما تكون ما تقيسه الفقرة أقل مما هو متوقع من المحتوى (Williams,1997) .

ويذكر كاميلي وشيبارد بأنه بالرغم من أن مصطلح الأداء التفاضلي للفقرة يستخدم كمصطلح مرادف لمصطلح تحيز الفقرة، إلا أن الأداء التفاضلي يستخدم في الكشف عن الفروق بين المجموعات متعددة الأبعاد، لتحديد التحيز بشكل غير مباشر، وذلك عند تساوي الأفراد على القدرة الأولية، في كلا
المجموعتين واختلافهم على القدرة الثانوية (Camili & Shepard, 1994) .

ويعتبر الأداء التفاضلي شرطا ضروريا، لكنه غير كاف للحكـم علـى تحيـز الفقرة، فإذا كانت الفقرة لا تظهر أداء تفاضلياً فأن الفقرة غير متحيـزة، أمـا فـي حالة وجوده فإن ذلك غير كاف من أجل الإعلان عن تحيز الفقرة، لكن ذلك يتطلـب تحليل إضافي لتحيز الفقرة(على سبيل المثال تحليل المحتوى، تقييم تجريبي) من أجل التحقق من تحيز الفقرة. (Zumbo ,1999) .

4.2 مصادر التحيز:

هناك العديد من مصادر التحيز المختلفة والتي تعد سبباً رئيسياً لظهور التحيز في الاختبارات النفسية والتربوية ومن أهم هذه المصادر:

1- الجنس: الفروق بين الجنسين في القدرات العقلية المقاسة هـي ظـاهرة قليلـة الحجم، وأن الفروق بين الجنسين في بعض خصائص الفقرات هي فروق غيـر مرتبطة بفروق ذات دلالة بالعامل العام .

2- العرق: ارتبط هذا التحيز باستخدام الاختبارات العقلية مع أفراد من مجموعات عرقية مختلفة من الأقليات (Jensen, 1980) .

ويرى ثورندايك أن فروقا في متوسط الأداء بين الجنسين قد تظهـر لـصالح الذكور في بعض الاختبارات ذات الطبيعة الذكورية، ولصالح الإناث في الاختبارات ذات الطبيعة الأنثوية، أما إذا ظهرت في الاختبارات ذات المحتوى المحايد ، فربما يعود ذلك إلى متغيرات بيئية وواقعية، يسهل تحديدها ولكـن يـصعب قياسـها أو ضبطها أو تقييم تأثيرها Thorndike, 1982**)**.

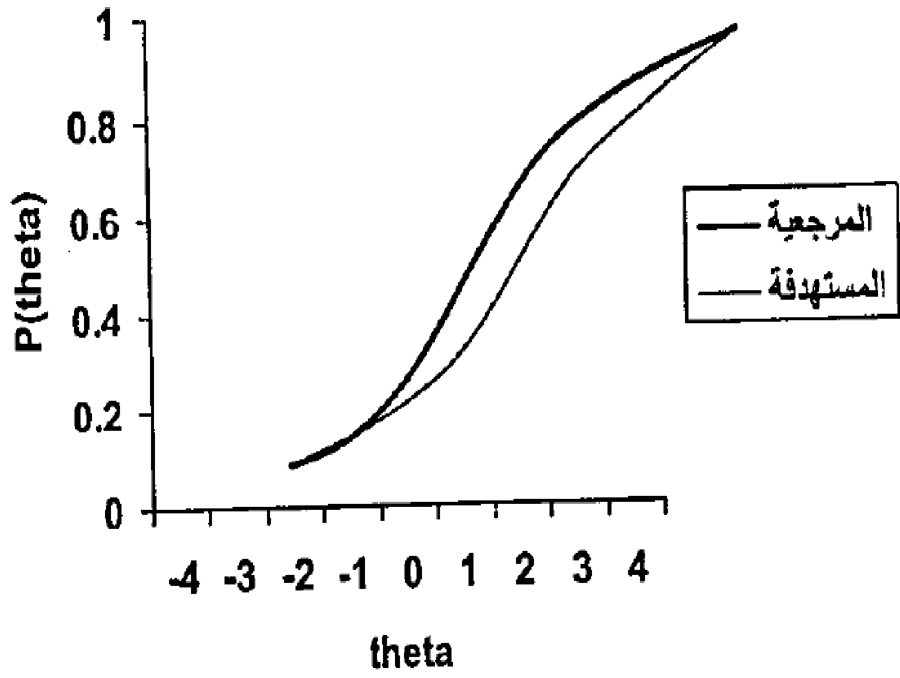
3- اختلاف المستويات الاقتصادية والاجتماعية يؤدي إلى فـروق ثقافيـة بـين المجموعات الفرعية المختلفة في المجتمع وبالتالي يؤدي إلـى تحيـز ثقـافي (Jensen, 1980)**.**

ويمكن القول أن الفقرة تظهر أداء تفاضلياً عندما تؤدي وظيفتها بشكل مختلف في المجموعات المختلفة عند تساوي مستوى القدرة لأفرادها حيث تكون الفقرة في إحدى المجموعتين أكثر أو اقل صعوبة مقارنة مع المجموعات الفرعية الأخرى، بعد ضبط مستوى القدرة لكلا المجموعتين، وهذا يعني أن الفقرة تعمل بشكل مختلف في كلا المجموعتين.

ويقسم الأداء التفاضلي للفقرة تبعا لطبيعة التفاعل بين مستوى القدرة وعضوية المجموعة إلى نوعين هما:

1- **الأداء التفاضلي المنتظم**: وهو الذي يظهر عندما ينتج تفاعل رتبي بين مستوى القدرة وعضويه المجموعة، أي أن احتمال الإجابة الـصحيحة للفقـرة يكـون للمجموعة A اكبر منه للمجموعة B عند جميع مستويات القدرة .

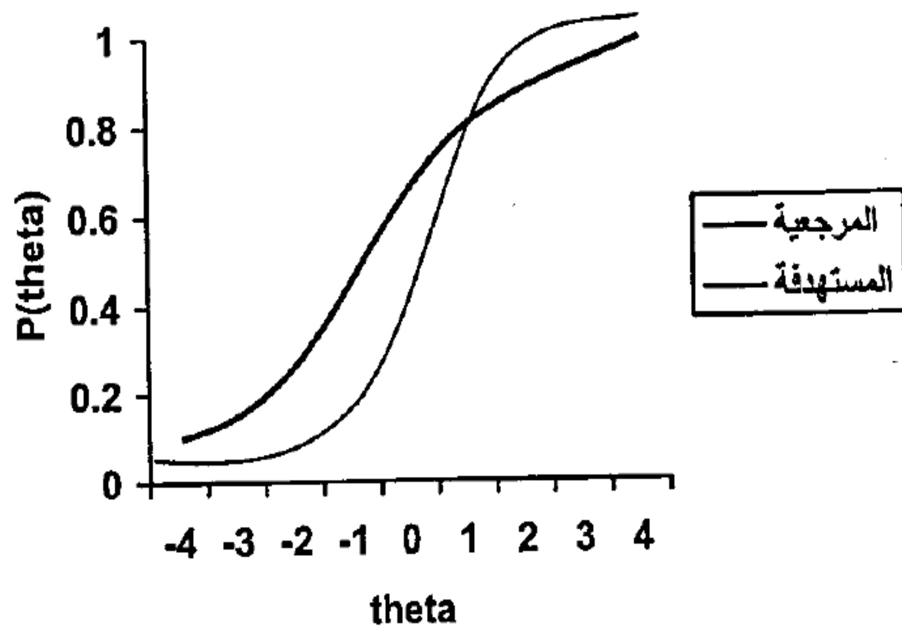
فكما هو مبين في الشكل(1) يلاحظ عدم وجود تفاعل بـين مـستوى القـدرة وعضوية المجموعة، وإنما احتمال الإجابة الصحيحة علـى الفقـرة دائمـاً أكبـر للمجموعة المرجعية.



الشكل(1)
الأداء التفاضلي المنتظم

2- **الأداء التفاضلي غير المنتظم (المتقاطع)**: وهو الذي يظهـر عنـدما يحـدث تفاعل غير رتبي بين مستوى القدرة وعضوية المجموعة أي أن فروق احتمـال الإجابة الصحيحة ليست ثابتة عند جميع مستويات القدرة .

والشكل(2) يبين ذلك، حيث يلاحـظ الأداء التفاضـلي لـصالح المجموعـة المستهدفة عند المستويات الدنيـا من القدرة، أما عنـد المـستويات العليـا فكـان لصالح المجموعة المرجعية، فعند تقاطع المنحنيين يكـون هنـاك تفاعـل بـين مستوى القدرة والمجموعـة، وبالتـالي ظهـور الأداء التفاضـلي غيـر المنـتظم (Swaminathan & Rogers, 1990)



الشكل(2)

الأداء التفاضلي غير المنتظم

5.2 طرق الكشف عن التحيز والأداء التفاضلي للفقرة:

لقد اقترح علماء القياس عدة طرق وأساليب إحصائية للكشف عن تحيز الفقرات في جميع أشكال الاختبارات، وتهدف هذه الطرق والأساليب الإحصائية لتنقية الاختبارات من التحيز ولتطبيق مبدأ العدالة والمساواة بين المجموعات المختلفة (Camili & Shepard, 1994).

ومن أهم طرق الكشف عن الأداء التفاضلي:

1.طريقة تحليل التباين:

بقيت هذه الطريقة من أشهر الطرائق المستخدمة في الكشف عن تحيز الفقرات حتى نهاية الثمانينات، وللكشف عن التحيز في هذه الطريقة، يتم تطبيق الاختبار على مجموعتين فرعيتين أو أكثر، حسب تصميم ثنائي، تتمثل متغيراته المستقلة بالفقرات وعضوية الأفراد في المجموعات الفرعية، بينما المتغير التابع هو عدد الأفراد الذين يستجيبون للفقرة بشكل صحيح، وينظر للتفاعل الذي تتحقق دلالته

الإحصائية بين المجموعات الفرعية والفقرات كمؤشر على وجود التحيز أو عدمـه (Osterlind, 1983) .

ويتم التعبير عن المعادلة الخطية لمصادر التباين كالتالي:

$$\sigma^2 = \sigma^2_i + \sigma^2_j + \sigma^2_{ij} + \sigma^2_{k(ij)} \text{----------(1)}$$

حيث أن:

σ^2_i: التباين الناتج عن صعوبة الفقرة.

σ^2_j: التباين الناتج عن اختلاف القدرة بين المجموعات.

σ^2_{ij}: التباين الناتج عن التفاعل بين القدرة والمجموعة .

$\sigma^2_{k(ij)}$: تباين الخطأ.

ونستطيع الحكم على عدم وجود تحيز للفقرة لمجموعه ما، من خلال اختبـار الفرضية الصفرية والتي تشير إلى عدم وجود تفاعل بين الفقرات والمجموعة عنـد مستوى دلالة معين، وعند رفض الفرضية الصفرية وقبول الفرضـية البديلـة يـتم استخدام الاختبارات البعدية لتحديد الفقرات المتحيزة.

ومن الانتقادات الموجهة لهذه الطريقة أن التفاعل بين الفقرة والمجموعـة قـد يكون مصدره اختلاف القدرة بين المجموعـات ولـيس تحيـز الفقـرة(Camili & Shepard,1994)

2.طريقة الصعوبة المحولة:

وضع هذه الطريقة كل مـن أنجـوف وفـورد (Angoff & Ford, Cited in Croker & Algina, 1986) في عام 1973، وتشبه هذه الطريقة طريقة تحليل التباين فهي تنظر إلى التفاعل بين الفقرة والمجموعة كمؤشر على وجود التحيز، بمعنى أن الفقرة تكون أكثر صعوبة في إحدى المجموعات مقارنة مع المجموعات الأخـرى، إلا أنها تختلف عنها في تحديد الفقرات المتحيزة من خلال درجة انحرافها عن الخط البياني الذي يظهر عدم تفاعل بين الفقرة والمجموعـة، ويمثل هـذا الخـط عـدم التحيز، ويتم الكشـف عن هذا التحيز فـي هـذه الطريقة حسـب الخطوات التالية (Camilli & Shepared, 1994):

1- حساب معامل صعوبة الفقرة (P) لكل فقرة في كلا المجموعتين الفرعيتين .

2- تحويل معامل صعوبة الفقرة (P) إلى العلامة المعيارية الزائية(Z-score).

3- تحويل قيمة (Z) إلى القيمة المعيارية (Δ)، حسب المعادلة: $\Delta = 4Z+13$

4-تمثيل قيم (Δ) بيانياً بحيث يوضع على المحور السيني قيم Δ للمجموعة الفرعية الأولى، وعلى المحور الصادي قيم Δ للمجموعة الفرعية الثانية.

5- يرسم خط الانحدار لقيم Δ للفقرات المحسوبة .

6- يحسب مؤشر التحيز وهو بعد النقطة عن المحور الرئيسي حسب المعادلة:

$$di = \frac{Bxi+|A-yi|}{\sqrt{B^2+1}} \quad \text{-----(2)}$$

حيث:

di: المسافة بين النقطة والمحور الرئيسي.

Xi: قيم Δ في المجموعة X.

Yi: قيم Δ في المجموعة Y.

A: المقطع الصادي لخط الانحدار.

B: ميل خط الانحدار.

فكلما زادت المسافة di كانت الفقرة أكثر تحيزاً، ويشير استرلند(Osterlind) إلى أن الفقرة تكون متحيزة إذا زادت المسافة بين النقطة والمحور الرئيسي عن انحراف معياري واحد (Osterlind,1983).

ومن أهم الانتقادات الموجهة لهذه الطريقة أنها تفترض تساوي المجموعات في القدرة، وتعمل على خلق توزيعين متماثلين في المجموعات من خـلال التحويـل، ويمثل ابتعاد النقاط عن خط الانحدار مشكلةً أخرى، حيث تحول المسافة من علاقـة غير خطية إلى علاقة خطية(Hoover & Kolen,1984).

3.طريقة تمييز الفقرة

في هـذه الطريقة يسـتخدم معامـل ارتباط الفقـرة مـع الـدرجة الكلية للاختبـار(Biserial Correlation Point) كمؤشر على قدرتها التمييزية، ويستخدم معامل التمييز في اتخاذ قرار بحذف الفقرة أو إبقائهـا فـي الاختبـار، حيـث أن الفقـرة ذات الارتباط الضعيف مع الاختبار لا تساهم في إضافة تباين حقيقي إلـى

تباين الاختبار، ويحسب مؤشر التمييز للفقرة لدى مجموعة من الأفراد وفق المعادلة التالية (Crocker & Algina, 1986):

$$r_{pbis}= \sqrt{\frac{p_i}{1-p_i}}.\frac{u_+ - u_\chi}{\sigma_\chi} \quad ------(3)$$

حيث u_+: الوسط الحسابي لعلامات جميع الأفراد الذين أجابوا على الفقرة إجابة صحيحة.

u_χ: الوسط الحسابي لعلامات جميع الأفراد الذين استجابوا للفقرة .

σ_χ: الإنحراف المعياري لعلامات المجموعة

p_i : صعوبة الفقرة .

ويتم ترتيب معاملات الارتباط تصاعديا أو تنازليا، والفقرة التي يكون معامل ارتباطها مع الاختبار في النصف العلوي للمجموعة الأولى وفي النصف السفلي للمجموعة الثانية يمكن تصنيفها على أنها فقرة متحيزة، لذا تعتبر الفقرة متحيزة إذا اختلف معامل تمييزها في المجموعة الأولى بدرجة ملحوظة عن المجموعة الثانية(Ironson, 1979).

4.طريقة كاي تربيع:

وضع هذه الطريقة فيشر وتعتبر هذه الطريقة أن الفقرة متحيزة، إذا اختلفت احتمالات الإجابة الصحيحة عنها من قبل المفحوصين ذوي القدرة المتساوية من المجموعات المختلفة (المرجعية المستهدفه)، اختلافا أكثر مما هو متوقع بمحض الصدفة (Crocker & Algina, 1986).

وإيجاد قيمة مربع كاي كمؤشر لتحيز الفقرة يتطلب الإجراءات التالية:

1. إيجاد مستويات القدرة المختلفة من خلال تقسيم مدى العلامات الكلي للاختبار إلى فئات
2. تحديد عدد الأفراد في كل مجموعه من المجموعات ولكل فئة من الفئات.
3. إيجاد عدد الأفراد الذين أجابوا إجابة صحيحة عن الفقرة في كل فئة لكل مجموعة من المجموعات (المرجعية المستهدفة).

4. حساب القيم المتوقعة لكل مجموعة، ويتم فيه حساب قيمة مربع كاي من خلال المعادلة الإحصائية التالية:

$$\chi_s^{\ 2} = \sum_{j=1}^{J} \frac{(O_{1j} - P_{.j}N_{1j})^2}{P_{.j}N_{1j}} + \sum_{j=1}^{J} \frac{(O_{2j} - P_{.j}N_{2j})^2}{P_{.j}N_{2j}} \quad \text{-------------(4)}$$

حيث :

N_{2j},N_{1j}: عدد مفحوصي المجموعتين الأولى والثانية الذين لهم علامات في المستوى j .

O_{2j},O_{1j}: عدد مفحوصي المجموعتين الأولى والثانية الذين أجابوا إجابة صحيحة على الفقرة في المستوى j .

P_{1j}: نسبة مفحوصي المجموعة الأولى الذين أجابوا بشكل صحيح على الفقرة في المستوى j .

P_{2j}: نسبة مفحوصي المجموعة الثانية الذين أجابوا بشكل صحيح على الفقرة في المستوى j .

P_j: نسبة المفحوصين من المجموعتين الذين أجابوا إجابة صحيحة على الفقرة عند مستوى j .

O_{2j},O_{1j}: القيم الملاحظة .

P_jN_{2j}, P_jN_{1j}: القيم المتوقعة .

وللكشف عن التحيز يتم مقارنة قيمة(χ^2)المحسوبة مع قيمة(χ^2)الحرجة عند درجة حرية $(K - 1)(j - 1)$

حيث:

K: تمثل عدد المجموعات.

j : تمثل عدد مستويات القدرة.

فإذا كانت قيمة (χ^2) دالة إحصائياً، دل ذلك على تحيز الفقرة، وكلما زادت قيمة كاي تربيع بعد أن تكون دالة إحصائياً تكون مؤشراً أكبر على تحيز الفقرة (Sheuneman, 1979).

5.طريقة تحليل المموهات:

تستخدم هذه الطريقة في أسئلة الاختيار من متعدد وتكون الأسئلة ذات 4-6 بدائل أحدها صحيحة، ويطلق على بدائل الإجابات غيـر الـصحيحة المموهـات، ويحاول مصمموا الاختبارات أن تكون المموهات متساوية في الجذب للمفحوصـين الذين لا يختارون البديل الصحيح، ويلاحظ أن مموهات إحدى الفقرات قد تختلف في جذبها للمفحوصين، حيث تختلق النسبة المئوية للمفحوصين الذين يختارون كل مموه من المموهات .

وتستخدم هذه الطريقة للكشف فيما إذا كان هناك جذب تمييزي نحو مموهـات السؤال بين المجموعتين ولذلك نجري لكل فقره اختبار الدلالة الإحصائية للفرضـية الصفرية (لا تختلف المجموعتان في التكرارات النسبية لاختيار المموهات). ورفض هذه الفرضية الصفرية عند مستوى دلالة معين يدل على ان هذه الفقرة متحيزة.

إن الاختبار الإحصائي المستخدم في هذه الطريقة هو اختبار مربع كاي، إذ يتم عمل مصفوفة لاستجابات الأفراد في المجموعتين على بدائل (مموهات) الفقرة، ثـم بعد ذلك نجري اختبار الدلالة الإحصائي مربع كاي في حالة كل مموه. ويستدل على وجود التحيز عندما تزيد قيمه مربع كاي عند مستوى دلالة محددة على قيمة معينـه أي في حالة اختيار المموه بشكل غير متساوي مـن قبـل المجموعـات الفرعيـة (Osterlind, 1983).

6. طريقة مانتل–هانزل (Mantel-Haenszel Method(MH)):

أستحدثت هذه الطريقة من قبل مانتل وهانزل في عام (1959)، ثـم طـورت هذه الطريقة كطريقة للكشف عن تحيز الفقرة في أواخر الثمانينيـات مـن القـرن الماضي من قبل هولاند وثاير وبالتحديد في عام 1988 (yan Song, 2005).

وتعتبر طريقة مانتل هانزل من أشهر الطرق وأكثرها استخداماً للكشف عـن الأداء التفاضلي، وتمتاز هذه الطريقة بسهولة حساباتها وإجراءاتها وفي هذه الطريقة تتم مقارنة الأداء على الفقرة بين مجموعتين فرعيتين إحداهما تـسمى المجموعـة المرجعية (الأغلبية) .

وتسمى الأخرى المجموعة المستهدفة(الأقلية) (Nandakumer,1993).

ولتطبيق هذه الطريقة يتم تنظيم البيانات في مصفوفة ثنائية البعد(2×2) كمـا هو الحال في الجدول الآتي:

جدول(1)

مصفوفة ثنائية البعد/مانتل- هانزل

المجموعة	إجابة صحيحة	إجابة خاطئة	المجموع
المرجعية (r)	A_j	B_j	n_{Rj}
المستهدفة(f)	C_j	D_j	n_{Fj}
المجموع	M_{1j}	M_{0j}	T_j

بحيث تشير الرموز الواردة في الجدول السابق إلى ما يلي:

1: الإجابة صحيحة، 0 : الإجابة خاطئة، (j): مستوى القدرة

n_{Rj} : عدد أفراد المجموعة المرجعية.

n_F: عدد أفراد المجموعة والمستهدفة.

A_j : عدد الأفراد في المجموعة المرجعية الذين أجابوا عن الفقرة إجابة صحيحة.

C_j: عدد الأفراد في المجموعة المستهدفة الذين أجابوا عن الفقرة إجابة صحيحة.

B_j : عدد الأفراد في المجموعة المرجعية الذين أجابوا عن الفقرة إجابة خاطئة.

D_j: عدد الأفراد في المجموعة المستهدفة الذين أجابوا عن الفقرة إجابة خاطئة.

M_{1j}: عدد أفراد المجموعتين الذين أجابوا عن الفقرة إجابة صحيحة.

M_{0j}: عدد أفراد المجموعتين الذين أجابوا عن الفقرة إجابة خاطئة.

يتم حصر البيانات وتنظيمها في جدول، وبعد ذلك يتم تطبيق معادلة مانتل هانزل (Mantel-Haenszel MH) $\chi 2$. على النحو الآتي:

$$MH_{\chi^2} = \frac{\left(\left| \sum_{J=1}^{S} \left[A_J - E\left(A_J\right) \right] \right| - 0.5 \right)^2}{\sum_{J=1}^{S} Var\left(A_J\right)} \quad ----(5)$$

حيث:

$E(A_j)$: القيمة المتوقعة لعدد الإجابات الصحيحة لأفراد مجموعة الأغلبية.

$Var(A_j)$: تباين A_j .

ولإيجاد القيمة المتوقعة يتم تطبيق المعادلة الآتية:

$$E(A_J) = \frac{(n_{rj} \cdot m_{1j})}{T_J} \quad \text{-----------(6)}$$

كما ويمكن إيجاد قيمة $Var(A_j)$ من خلال المعادلة:

$$Var(A_J) = \frac{n_{rj} n_{fj} m_{1j} m_{0j}}{T^2_J (T_J - 1)} \quad \text{-----------(7)}$$

فإذا كانت قيمة $\chi^2 MH$ دالة إحصائيا دل ذلك على أن هنالك تحيزا للفقرة عبر المجموعات.

وهناك مؤشر آخر يسمى نسبة الأفضلية (αMH) للدلالة على تحيز الفقـرة ويتم إيجاد قيمته باستخدام المعادلة التالية:

$$\propto_{MH} = \frac{\sum_{J=1}^{S} A_J D_J / T_J}{B_J C_J / T_J} \quad \text{-----------(8)}$$

فإذا كانت (αMH=1) فالفقرة لا تظهر أداء تفاضلياً، وإن كانـت (αMH<1) فهي تظهر أداء تفاضلياً لصالح المجموعة المـستهدفة، أمـا إذا كانـت(αMH>1) فالفقرة تظهر أداء تفاضلياً لصالح المجموعة المرجعية (Nandakumer,1993).

ويمكن استخدام طريقة مانتل-هانزل للكشف عن الأداء التفاضلي في الفقرات المتعددة الاستجابة، وقد تم تفضيل هذه الطريقة فـي هـذه الدراسـة لان فقـرات الاختبارات مقسمة إلى فقرات ذات إجابة منتقاة (اختيار من متعـدد) وفقـرات ذات إجابة مصاغة(مقالية) ثنائية الإجابة، وكذلك لأنها تستخدم للعينات الصغيرة والكبيرة.

7. طريقة الانحدار اللوجستي:

تستخدم هذه الطريقة في الكشف عن الأداء التفاضلي المنتظم وغير المنـتظم (Swaminathan & Rogers, 1990). ويتم الكشف عن الأداء التفاضلي للفقـرة مـن خلال حسن المطابقة مع النموذج، عندما نضيف للنمـوذج متغيـر العـضوية فـي المجموعة ومتغير التفاعل بين العضوية في المجموعة والقدرة.

ومعادلة الانحدار اللوجستي التي نحدد بواسطتها احتمالية الاستجابة الصحيحة عن الفقرة:

$$\pi(u = 1) = \frac{e^z}{1 + e^z} \quad \text{--------}(9)$$

$$\ln\left[\frac{\pi ij}{(1 + \pi ij)}\right] Z = logit(\pi ij) \quad \text{--------}(10)$$

وعند حل معادلة الانحدار اللوجستي والتحويل إلى وحدة اللوجيت يتكون لدينا ثلاث معادلات (نماذج):

1- $Z=\beta 0+\beta 1\theta$

2 - $Z=\beta 0+\beta 1\theta+\beta 2G$

3- $Z= \beta 0+\beta 1\theta+\beta 2 G+\beta 3(\theta G)$

حيث أن:

G : العضوية للمجموعة.

θ : قدرة المفحوص المقدرة من خلال علامته على الاختبار.

β0 : المعلمة المشتركة

β1θ : الإختلاف بين المجموعات ناتج عن الاختلاف في قدرة المفحوصين وليس بسبب وجود الأداء التفاضلي.

β2G : وجود أداء تفاضلي منتظم بسبب الاختلاف في العضوية للمجموعة.

β3(θG) : وجود أداء تفاضلي غير منتظم بسبب التفاعل بين القدرة والعضوية في المجموعة.

يتم الاستدلال على وجود الأداء التفاضلي (المنتظم وغير المنـتظم) أو عـدم وجوده من خلال مقارنة النماذج الثلاث السابقة كما يلي:

1. يتم فحص النموذج الأول مع النموذج الثالث عند درجة حرية (2) باستخدام اختبار نسبة الأرجحية العظمى، فإذا لم تظهر فروق ذات دلالة إحصائية فإن ذلك يدل على أن الفقرة لا تبدي أي أداء تفاضلي .أما إذا ظهرت فروق ذات دلالة إحصائية، فإن ذلك يدال على أن الفقرة تبدي أداء تفاضلياً، وبالتالي ننتقل إلى الخطوة الثانية لتحديد نوع الأداء التفاضلي.

2. يتم فحص النموذج الثاني مع النموذج الثالث عند درجة حرية (1) باستخدام اختبار نسبة الأرجحية العظمى فإذا ظهرت فروق ذات دلالة إحصائية فإن ذلك يدال على أن الفقرة تبدي أداء تفاضلياً غير منتظم، وإذا لم تظهر فروق ذات دلالة إحصائية ننتقل إلى الخطوة الثالثة.

3. يتم فحص النموذج الأول مع النموذج الثاني عند درجة حرية (1) باستخدام اختبار نسبة الأرجحية العظمى، فإذا ظهرت فروق ذات دلالة إحصائية فإن ذلك يدال على أن الفقرة تبدي أداء تفاضلياً منتظماً (Hdalgo & Lopez-Pina, 2002).

8. طريقة منحنى خصائص الفقرة:

انبثقت هذه الطريقة للكشف عن الفقرات المتحيزة عن نظرية استجابة الفقرة، ويمثل المنحنى العلاقة الرياضية التي تربط بين احتمالية إجابة الفرد إجابة صحيحة على الفقرة i وبين القدرة أو السمة التي تقيسها فقرات الاختبار الذي يحتوي على الفقرة i، وهي دالة انحدار غير خطية، وبمعرفة درجات الاختبار للمفحوصين عند كل مستوى قدرة محدد يمكن رسم منحنى خصائص أي فقرة من فقرات ذلك الاختبار والذي يمثل خط الانحدار الذي يمر بمتوسط التوزيعات الشرطية لكل مستوى قدرة (Hambleton & Swaminathan, 1991)

ويتم تقدير الأداء التفاضلي للفقرة باستخدام منحنى خصائص الفقرة بعد معادلة مقياس القدرة، أي وضع المجموعتين على نفس تدريج القدرة، حيث يلاحظ عندها اختلاف في احتمال الإجابة عن الفقرة في كل مجموعة، ومن ثم استخلاص الأداء التفاضلي للفقرة من خلال الاختلاف في معالمهما (الصعوبة، التمييز والتخمين) وفق النموذج المعتمد. حيث تبين هذه المعالم مدى التشابه أو التباين بين منحنيات

خصائص الفقرة في المجموعات الفرعية، ومن ثم إيجاد الإحصائي الذي يعد مؤشراً للأداء التفاضلي للفقرة حسب معيار الطريقة المعنية.

وقد أورد كاميلي وشيبارد عدة طرق لقياس حجـم الأداء التفاضـلي للفقـرة وأهمها (Camili & Shepard, 1994):

أ - مؤشر المساحة: أن مقدار المساحة بين منحنيين من منحنيات خصائص الفقـرة يعد مؤشراً لمدى اختلاف المجموعتين من حيث أداؤهما في فقرة ما.وعليه فإن الفقرات ذات المساحة الكبرى بـين المنحنيـين تظهـر أداء تفاضـلياً بـين المجموعتين أكثر من فقرة بمساحة أصغر بين المنحنيين، ورغم أن مؤشرات المساحة لها ميزة إمكانية تمثيلها بيانياً وتصورها بسهولة إلا أن لها سلبية وهي أن التكاملات تعطي قيم لا نهائية،عندما تكون معالم التخمين لمجموعتين غير متساوية.

ب - فرق المعلم b: إن فرق المعلم b (معامل الصعوبة) عند ثبوت القدرة يعتبـر مؤشراً على الأداء التفاضلي للفقرة.وتشير القيمة الموجبة للمقدار (Δb) إلـى أن هناك أداء تفاضلياً لصالح المجموعة المرجعية، بينما القيمة السالبة للمقدار (Δb) تشير أن هناك أداء تفاضلياً لصالح المجموعة المستهدفة.

وللكشف عن الأداء التفاضلي للفقرة يتم الحكم على الفقرة بأنها لا تظهر أداء تفاضلياً إذا كانت منحنيات خصائص الفقرة متماثلة في المجموعـات المختلفـة، أي تكون الفقرة متساوية المعالم لأفراد المجموعات الذين لهم نفس المستوى من القـدرة (Croker and Algina,1986).

9. طريقة نسبة الأرجحية العظمى:

تستخدم هذه الطريقة للكشف عن الأداء التفاضلي في الفقرات ثنائية التـدريج والفقرات متعددة التدريج، ويستخدم الاختبار الإحصائي (Likelihood-Ratio Test) لفحص الفرضية الصفرية التي تنص على عدم وجود فروق ذات دلالة إحصائية في معالم الفقرات بين المجموعات أو عدم وجود فروق ذات دلالة إحصائية فـي دالـة استجابة الفقرة بين مجموعة وأخرى،.

ويتكون الاختبار الإحـصائي (Likelihood-Ratio Test) مـن الفـرق بـين نموذجين (Compact & Augmented) كالتالي:

$$Gj2 = -2Log\ L\ (Compact\ Model) + 2\ Log\ L(Augmented\ Model)$$

وهذا الإحصائي له توزيع مماثل لتوزيع مربع كاي والذي معادلته :

$$\chi^2_{LR} = -2\sum_{ij} f_{ij} \ln (E_{ij}/f_{ij}) \quad \text{-----------(11)}$$

ونموذج (Compact Model) قائم على افتراض أن المعـالم فـي المجموعـة المرجعية والمجموعة المستهدفة متساوية، فعندما يتم مطابقة البيانـات مـع معـالم منحنيات الاستجابة للفقرات تضبط بحيث تكـون متساوية فـي كـلا المجمـوعتين لكل الفقرات، ويتم اختيار التوزيع الطبيعي الـذي متوسـطة = صـفر وانحرافـة المعياري =1، *N(0,1)*، كسلم لمعالم الفقرات فـي المجموعـة المرجعيـة،أما فـي المجموعة المستهدفة فيتم تقدير المتوسط الحسابي والانحراف المعيـاري للتوزيـع الطبيعي للمجتمع بشكل متزامن مع معالم الفقرة.

أما النموذج (Augmented Model) فهو قائم على افتراض أن جميع المعـالم في المجموعة المرجعية والمجموعة المستهدفة متساوية باسـتثناء الفقـرة موضـع الدراسة فيتم تقدير المعالم لها بشكل مختلف في المجموعتين، ومـرة أخـرى يـتم اختيار التوزيع الطبيعي الذي متوسطة = صفر وانحرافه=1 *N(0,1)* كـسلم لمعـالم الفقرات في المجموعة المرجعية، أما في المجموعة المستهدفة فيتم تقدير المتوسـط الحسابي والانحراف المعياري للتوزيع الطبيعي للمجتمع بشكل متزامن مـع معـالم الفقرة (Kim & Cohen,1995).

6.2 الدراسات السابقة:

بعد مراجعة الأدب التربوي، والإطلاع علـى طـرق الكـشف عـن الأداء التفاضلي، تم عرض بعض الدراسات التي تناولت موضوع التحيز والكـشف عـن الأداء التفاضلي لمتغير الجنس ولمتغيرات أخرى في اختبـارات متنوعـة حـسب الترتيب الزمني من الأقدم إلى الأحدث.

أجرى كل من سكاجس ولزتس(skaggs and lists ,1992) دراسة هدفت إلـى المقارنة بين عدد من طرق الكشف عن تحيز الفقرة تبعا لاختلاف الجنس والعـرق وهي: طريقة المساحة بين منحنيي خصائص الفقرة، وطريقة كاي تربيع، وطريقـة مانتل هانزل (القيم المحولة إلى قيم دلتا)، وطريقة لورد كاي تربيع وطريقه مربـع الفرق بين احتمالية الإجابة الصحيحة للمجموعتين، وطريقة نظريه الاستجابة للمفردة المعتمدة على مجموع المربعات للفروق بين المعالم المختلفة، وطريقة مانتل هانزل المعتمدة على كاي تربيع، وكانت أداة الدراسة عبارة عن اختبار تحصيلي في مـادة الرياضيات مكون من (96) فقرة تم تطبيقه مرتين بفاصل زمني مدته (12) شـهرا وقد بلغ حجم العينة في التطبيق الأول (1300) فرد وفي التطبيق الثاني (400) فرد، حيث أظهرت نتائج الدراسة ان طريقة نظرية الاستجابة للفقرة والتي تعتمـد علـى مجموع المربعات هي أكثر الطرق فاعلية في الكشف عن تحيز الفقرات تلتها طريقة مانتل هانزل (القيم المحولة إلى قيم دلتا) وثم طريقة مانتل هانزل المعتمدة على كاي تربيع، حيث كانت طريقة كاي تربيع اضعف الطرق مقارنه مع الطرق الأخرى.

وأجرى كل من راجو دراسجو وسـيلند (Raju , Drasgow and Slind,1993) دراسة هدفت إلى المقارنة بين عدد من طرق الكشف عن تحيـز الفقـرات وهـي: طريقة المساحة بين منحنيي خصائص الفقرة، طريقة مانتل هانزل (قيم دلتا) وطريقة كاي تربيع، حيث استخدم الباحث اختبار لفظي للفئة العمرية (10-12) سنه مكـون من (45) فقرة، وتم تطبيق الاختبار على عينه مكونه مـن (839) طالبـا وطالبـة موزعين في عدد من المجموعات الفرعية المتعلقة بالجنس والعرق، وأظهرت نتائج الدراسة بأن طريقة مانتل هانزل (قيم دلتا) استطاعت الكشف عن الفقرات المتحيزة بنسبة اكبر من الطريقتين الأخيرتين، كما أظهرت نتائج الدراسة وجود درجة عالية من الارتباط بين كل من طريقتي: المساحة بين منحنيي خصائص الفقـرة وطريـق كاي تربيع حيث بلغ معامل الارتباط بين الطريقتين (0.92). بمعنـى أن الفقـرات المتحيزة بموجب طريقة المساحة بين منحنيي خصائص الفقرة هي أيـضا متحيـزة بموجب طريقة كاي تربيع.

وأجرى نينغ ولان (Ning&Lane,1996) دراسة هدفت إلى الكشف عن الأداء التفاضلي لمتغير الجنس في تقيم أداء الرياضـيات. وجـرى تحديـد المحتـوى أو العمليات الإدراكية (كالإختلافات في توزيعات القدرة)، التي قد تكـون ذات صـله بالأداء التفاضلي للفقرة، بالإضافة إلى فحـص ملائمـة الفقـرات لـشروط الأداء التفاضلي للفقرة وفقا لمتغير الجنس، و تكونت أداة تقييم الإدراك الكمي من 36 فقره متعددة التدريج حيث قسمت إلى أربعة نماذج، وكل نموذج يحتوي على تسع فقرات مختلفة، وكانت النماذج الأربعة (أ، ب، ج، د) قد وزعت عشوائيا ضمن كل صـف من صفوف السادس والسابع في مدارس مشاركه في مشروع الفهم الكمي في ولاية فيرجينليا الأمريكية، ودرجت إجابة كل طالب على كل فقرة من قبـل اثنـين مـن المدرجين أو المقدرين، وقد تألفت عينة الدراسة من (1782) طالباً وطالبـة، وقـد بينت النتائج أن (31) فقره من (33) فقره (94%من الفقرات) لم تظهر أداء تفاضليا ذا دلالة إحصائية في أداة التقييم الكمي. كما كانت نسب التوافق بين الإجرائين فـي كشف الأداء التفاضلي للفقرة عالية نسبيا، حيث تراوحت من 88% إلى 97% عند مستوى دلاله 0.05 .

وفي دراسة أجراهـا كـل من هامبلتـون ورانـد وسـنو (Hambleton, Snow, 1998 & Rand) هدفت للكشف عن التحيز لمتغير الجنس في اختبار للعلـوم بشكليه اختيار من متعدد واستجابة مفتوحة. طبق على عينة من (2000) طالب على مستوى الصف الثامن والصف العاشر و الصف الثاني عشر، يحوي (25) فقرة من نوع الاختيـار من متعدد و(25) فقرة متعددة الاستجابة، وأظهرت نتائج الدراسة أن عدد الفقرات المتحيزة في اختبار العلوم (17) فقرة مـن أصـل (25) لـصالح الذكور؛ وعدد الفقرات المتحيزة بطريقة دالة التحليل التمييزي اللوجستي (19) فقرة من أصل (25) لصالح الذكور.

وأجرت هاميلتون (Hamilton,1999) دراسة هدفت إلى استطلاع طرق الكشف عن الأداء التفاضلي لمتغير الجنس على فقرات اختبار العلوم متعدد التدريج لطلبـة الصف الثاني عشر، والذي أعطي كجزء من الدراسة الطولية التي يـشرف عليهـا المركز الوطني للإحصائيات التربوية (NELS) للتعليم للعام 1988. وقـد انـصب

التركيز على تحديد خصائص الفقرات التي يبدو فيها فروق كبيرة على نحو خاص. تابعت الدراسة عينة من طلبة الصف الثامن إلى الصفوف العاشر والثـاني عـشر باستخدام اختبارات معرفيه في أربعة مباحث، وأجرى المركز الوطني للإحصاءات التربوية دراسة لفاعلية المدارس الثانوية كدراسة إضافية مكملة، جرى فيهـا اخـذ عينة من 247 من طلبة الصف العاشر عام 1990 وجرت متابعتهم حتـى الـصف الثاني عشر، وقد اشتمل اختبار العلوم من نوع الاختيار من متعدد للـصف الثـاني عشر على 52 فقرة وكانت مدة الاختبار 20 دقيقة. كذلك قامت عينة فرعية مكونـه من 2204 من طلبة الصف الثاني عشر لدراسة فاعلية المدرسة الثانوية بالإجابة عن 4 فقرات متعددة التدرج، وكانت المدة المحددة لكل فقره 10دقائق، وكان يطلب فيها من الطلبة كتابة تفسير أو رسم توضيحي. وتم تقسيم كل فقرة إلى عـدة خطـوات، صحح كل منها باستخدام فئات للإجابات المحتملـة. وبعـد الانتهـاء مـن عمليـة التصحيح، صمم نظام لتجميع الدرجات في فئات منتظمة لكل فقـرة، وقـد كـشفت الدراسة انه هناك فقرة واحدة كان فيها أداء تفاضلياً لصالح الذكور، وكانت النتـائج مماثلة لتلك التي تم الحصول عليها في الاختبار ثنائي التدريج. وفي كلتا الحـالتين، كانت فروق الجنس اكبر ما تكون في الفقرات التي يدخل فيها التـصور البـصري وتتطلب استحضار المعرفة المكتسبة خارج نطاق المدرسة .

وأجرى كاثـرين وشيـووان (Katherine & Shaun,2001) دراسـة تهـدف إلـى الكشـف عن الفقـرات ذات الأداء التفاضلي لمتغير الجـنس وتـأثير تغيـر مواقعهـا في نمـوذج اختبار مستـوى فـي الرياضـيات (MMPE Midwestren Mathematics Placement Exam)، يستخدم لتحديد المساقات التي سيدرسها الطالب المستجد عند قبوله في الجامعة في تخصص الرياضيات، ويتكون الاختبار من (40) فقرة من نوع الاختيار من متعدد، ويقيس المفاهيم الرياضية في الجبر، والمثلثـات، والهندسة، وتتوزع فقرات الاختبار على النحو الآتي: (18) فقرة تقيس مفاهيم الجبر، (12) فقرة في موضوع المثلثات، (10) فقرات في موضوع الهندسة، ويعد الاختبار لجان متخصصة في تطوير اختبارات الرياضيات، وتضعه في نموذجين: النمـوذج الأول وزعت فيه الفقرات عشوائياً، في حين وزعت الفقرات في النموذج الثاني من

الأسهل إلى الأصعب، وطبق الاختبار على عينة مكونة مـن (1665,1066) طالبـاً وطالبةً على الترتيب، وكانت علامة المفحوص الكلية هي الإجابة الصحيحة علـى الفقرات على أن يجيب على الفقرات خلال (75) دقيقة، وأظهرت نتائج الدراسة أن فقرات موضوع الجبر في النموذج الثاني كانت أسهل على الإناث من الذكور منهـا في النموذج الأول، في حين أن الفقرات في موضوعي المثلثات والهندسـة كانـت أسهل عند الذكور منها على الإناث في كلا النموذجين .

وأجرى جرادات(2003) دراسة هدفت إلى المقارنة بين طريقة مانتل هانزل وطريقة الصعوبة المحولة للكشف عن تحيز الفقرات، استخدم فيها بيانـات تقنـين اختبار كاليفورنيا للتفكير الناقد للبيئة الأردنية، طبقه على عينة مكونة مـن(1485) طالبا وطالبة من جامعة مؤتة، (789 أنثى)، (696 ذكر)، وتكون الاختبار من (34) فقرة وأظهرت نتائجه: أن طريقة معامل الصعوبة أكثر قدرة على كـشف الفقـرات المتحيزة للجنس وتماثل الطريقتين في كشف الفقرات المتحيزة للحقـل الأكـاديمي، واتفاق الطريقتين على اعتبار الفقرة متحيزة للجنس.

وأجرى البستنجي(2004) دراسة هدفت إلى مقارنـة الأداء التفاضـلي بـين طريقتين من طرق النظرية الكلاسيكية للكشف عن تحيز الفقرات لمتغير الجـ ـنس (طريقة مانتل هانزل وطريقة الانحدار اللوجستي)، وطريقتين من طرق النظريـ ـة الحديثة في القياس (طريقة مؤشر الفرق في المساحة لمنحنى خصائص الفقرة لمتغير (ثنائي المعلم وثلاثي المعلم) واستخدم اختبار يقيس ثلاث قدرات عقلية خاصة لأفراد أردنيين من فئة (15-16) سنة وطبق الاختبار على عينة مكونه من (950) طالبـا وطالبة منهم (466) ذكور (484) إناث، وتوصلت الدراسة إلى اتفاق طريقة مؤشر الفرق في المساحة لنموذج ثنائي المعلم مع طريقة فرق المـساحة لنمـوذج ثلاثـي المعلم في فقرة واحدة في اختبار القدرة اللفظيـة وفقـرتين فـي اختبـار التفكيـر الرياضي، وثلاث فقرات في اختبار القدرة المكانية، وعدم اتفاق الطريقتين فـي أي فقرة لاختبار المفردات بالتضاد وكان التحيز لصالح الذكور في اختبـار التفكيـ ـر الرياضـي، والقدرة المكانية، ولصالح الإناث في اختبار القدرة اللفظية واخـتلاف

التحيز كان أكبر ما يمكن في المستويات العليا من القدرة وأقل مـا يمكـن فـي المستويات الدنيا.

وأجرت (العطيوي، 2004) دراسة بهدف الكشف عن الأداء التفاضلي لمتغير الجنس في اختبار تحصيلي لمادة العلوم مكون من (60) فقرة من نوع الاختيار من متعدد، بواقع (15) فقرة لكل من الأحياء والفيزياء والكيمياء وعلوم الأرض. طبـق على عينة مكونة من (110) طلاب مستخدمة طريق المساحة الفاصلة بين منحنيات خصائص الفقرة وطريقة مانتل هانزل في الكشف عن التحيز، وتوصلت إلى وجود (21) فقرة متحيزة لمتغير الجنس حسب طريقة منحنيات خصائص الفقرة منها (11) فقرة تحيز غير منتظم، (10) فقرات تحيز منتظم منها (6) فقرات لصالح الـذكور، (4) فقرات لصالح الإناث، و كشفت طريقة مانتل عن (11) فقرة متحيـزة لمتغيـر الجنس منها (7) فقرات لصالح الذكور و (4) لـصالح الإنـاث، وعـدد الفقـرات المشتركة بين الطريقتين (5) ومجموع الفقرات المتحيزة باتفاق الطـريقتين (27) فقرة.

وأجرى (الهريشات، 2004) دراسة هدفت إلى تحري التحيز في فقرات اختبار أيسر للتفكير الاستدلالي المقنن لطلبة الصف الأول الثانوي في الأردن مـن طلبـة المدارس الحكومية من خلال طريقتي مانتل هانزل وطريقة الصعوبة المحولة وذلك تبعاً لمتغيرات التالية: الجنس والتخصص الأكاديمي (علمي، أدبي) ومنطقة الـسكن (شمال، وسط، جنوب) حيث تكون الاختبار من (70) فقرة من نوع الاختيـار مـن متعدد وتم تطبيقه على عينة مكونة مـن (591) طالبـاً وطالبـة للعـام الدراسـي 2004/2003، وأظهرت نتائج الدراسة تحيز (5) فقرات للجنس منها:(3) لـصالح الذكور وفقرتان لصالح الإناث وذلك باستخدام طريقة مانتل- هـانزل، فـي حـين كشفت طريقة الصعوبة المحولة تحيز (8) فقرات منها: (6) فقرات لصالح للـذكور وفقرتان لصالح الإناث واتفقت الطريقتان على تحيز (5) فقرات، وكـشفت طريقـة مانتل- هانزل عن تحيز (7) فقرات للحقل الأكاديمي منها: (6) فقرات لصالح الفرع العلمي وفقرة واحدة لصالح الفرع الأدبي، في حين كشفت طريقة الصعوبة المحولة تحيز (10) فقرات منها: (9) فقرات لصالح الفرع العلمي وفقرة واحدة.

وأجرى (حمادنة، 2007) دراسة هدفت إلى مقارنة الأداء التفاضـلي لفقـرات تحديد الكفاءة اللغوية في اللغة الإنجليزية في الجامعـات الأردنيـة (اليرمـوك، آل البيت، العلوم والتكنولوجيا) تبعاً لمتغير الجنس وفرع الثانوية العامة (علمي، أدبي)، حسب طريقتي مانتل هانزل وطريقة فرق المساحة بين منحنيات خصائص الفقرة في النموذج اللوجستي ثلاثي المعلمة، والتعرف على درجة التوافق بين الطريقتين، وقد تم اختيار ثلاثة نماذج من اختبارات تحديد الكفاءة اللغوية فـي اللغـة الإنجليزيـة، وكانت جميع فقرات الاختبارات من نوع الاختيار من متعدد، وبلغ عدد الفقرات في الجامعات الثلاثة (50, 50 ,100) فقرة على الترتيب، وتقيس الفقرات فـي مجملهـا مهارات مشتركة في اللغة الإنجليزية هي: القواعد والمفردات، الاستيعاب المقروء، والكتابة، وتكونت عينة الدراسة من (1935) طالباً وطالبة، وقد أظهرت النتـائج أن أكثر المهارات التي أظهرت فقراتها أداء تفاضلياً لمتغيـر الجـنس هـي القواعـد والمفردات، وكانت في معظمها لصالح الإناث، وكانت نسب الفقرات التي أظهـرت أداء تفاضلياً لمتغير الجنس (24%, 21%, 3%) لجامعة اليرموك، جامعـة العلـوم والتكنولوجيا، وجامعة آل البيت على الترتيب،
لصالح الفرع الأدبي واتفقت الطريقتان على تحيز (7) فقرات.

وأجرى لي (Le,2009) دراسة هدفت إلى الكشف عن الأداء التفاضلي لمتغير الجنس عبر الدولة واللغة التي قدم بها الاختبار، في فقـرات اختبـار العلـوم فـي الدراسة التجريبية لاختبار(2005) PISA، ومقارنة الأداء التفاضلي لمتغير الجـنس عبر الدول المختلفة المشاركة بالاختبار، وعبر اللغات التي يقدم بها الاختبار للطلاب في مختاف الدول المشاركة في هذا الاختبار، وهدفت الدراسة أيضاً إلـى اسـتخدام خصائص الفقرات، والأبعاد الرئيسية للمعرفة العلمية(السياق، المعرفة، الكفايـات، الاتجاهات) في فقرات اختبار العلوم لاستكشاف أوجه القوة والضعف لأداء الإنـاث والذكور، وقد بلغ عدد الطلبة المشاركين في هذا الاختبار (83000) طالبا وطالبـة من (50) دولة، وبلغ عدد فقرات الاختبار (210) فقرات قدمت للطلاب فـي (60) لغة مختلفة. وتم استخدام طرق النظرية الحديثة(PCM)،(PLM) للكشف عـن الأداء التفاضلي لمتغير الجنس عند كل مجموعه لغة، وكل الـدول المـشاركة فـي هـذا

الاختبار، وقد بينت نتائج الدراسة اثر الدولة ولغة الاختبار علـى الأداء التفاضـلي لمتغير الجنس في فقرات اختبار العلوم، حيث أظهرت نتائج الدراسـة وجـود أداء تفاضلي لفقرات اختبار العلوم، يعزى لمتغير الجنس تبعا لاختلاف اللغة التي قدم بها الاختبار، وبنسبة عالية، كذلك وجود أداء تفاضلي في فقرات اختبار العلوم لمتغيـر الجنس عبر الدول المختلفة المشاركة بالاختبار وبنسبه عالية أيضا، كذلك بينت نتائج الدراسة أثر الأبعاد وخصائص الفقرات، حيث فضل الذكور الفقرات ذات الاختيـار من متعدد والإجابة القصيرة، وبعدي السياق والكفايات، بينما فضلت الإناث الفقرات ذات الإجابة الطويلة، وبعد المعرفة العلمية، كما بينت الدراسة أن الكشف عن الأداء التفاضلي وفقاً لمتغير الجنس عبر الدولة واللغة ضـروري ويـسهم فـي تطـوير الاختبارات الدولية بدرجة كبيرة .

وأجرى (الطراونة، 2011) دراسة لتحري التحيز في فقرات الاختبار الوطني لطلبة الصف الثامن في مادتي الرياضيات والعلوم للعام 2010 والمعد مـن وزارة التربية والتعليم الأردنية وذلك باستخدام طريقة مانتل-هانزل، حيث تكون الاختبـار من(60) فقرة من نوع الاختيار من متعدد: منها(30) للعلوم و (30) للرياضـيات. وقد طبق الاختبار على عينة تتكون من (67455) طالبا وطالبـة، حيـث كـشفت الدراسة عن تحيز (9) فقرات لصالح الإناث في اختبـار العلـوم، و (4) فقـرات لصالح الإناث في اختبار الرياضيات، و تحيز (17) فقره لـصالح سـلطة التعلـيم الخاص في كلا الاختبارين، وكشفت الدراسة عن تحيز (7) فقرات لصالح الـشمال مقارنة بالجنوب وتحيز (3) فقرات لصالح الوسط مقارنة بالجنوب وتحيز (12) فقرة لصالح الشمال مقارنة بالجنوب.

وأجرى (البرصان، 2013) دراسة للكشف عن الأداء التفاضلي لمتغير الجنس لفقرات الاختبار الوطني الأردني لضبط نوعية التعليم لمـادة الرياضـيات للـصف العاشر، وذلك لكل من المستويات الأربعة التي تصنف وزارة التربية والتعليم طلبتها فيها، وهي المستوى المتقدم، والماهر، والماهر جزئياً، والمبتدئ. حيث بلـغ عـدد أفراد عينة الدراسة (43486) طالباً وطالبة اختيروا مـن طلبـة الـصف العاشـر الأساسي في الأردن بواقع (21330) طالباً و (22138) طالبة، ولتحقيـق أهـداف

الدراسة تم الحصول على استجابات الطلبة للاختبار الوطني الأردني لمادة الرياضيات المكون من (30) فقرة، ثم تم تحليلها لمستويات الطلبة الأربعة باستخدام طريقة مانتل هانزل للكشف عن الأداء التفاضلي، وكشفت نتائج الدراسة عن الفقرات التي تظهر أداء تفاضلياً في كل مستوى، وقد كان عدد الفقرات يزيد كلما اتجهنا من المستوى الأعلى للمستوى الأدنى، حيث تراوحت نسبة الفقرات التي تظهر أداء تفاضلياً بين 30% للمستوى المتقدم، و 67% للمستوى المبتدئ.

بعد الإطلاع على الدراسات السابقة يلاحظ أن أداء الذكور في معظم الدراسات أعلى من أداء الإناث في العلوم، والقدرة الرياضية، والتفكير الرياضي، والمثلثات والهندسة أما في القدرة اللفظية والقواعد والمفردات، والجبر فإن أداء الإناث كان أعلى.

أما بالنسبة لطرق الكشف عن الأداء التفاضلي لأي من الجنسين فلا يوجد هناك طريقة فعالة للكشف عن التحيز والأداء التفاضلي، والدليل على ذلك عدم الاتساق في نتائج هذه الدراسات ويرجع سبب ذلك إلى أن الأداء التفاضلي للفقرة يعتمد على عدة عوامل منها حجم العينة، ونوع الأداء التفاضلي (منتظم، غير منتظم)، وعدد فقرات الاختبار، والتطبيقات السيكومترية التي تختلف باختلاف النظرية الكلاسيكية والنظرية الحديثة .

كما يلاحظ أن كل الدراسات التي أجريت للكشف عن الأداء التفاضلي باستخدام طرق مختلفة كشفت عن وجود فقرات تظهر أداء تفاضلياً في فقرات الاختبارات المستخدمة مما يؤكد على ضرورة التنبه لمشكلة التحيز والأداء التفاضلي، والعمل قدر الإمكان على تقديم اختبارات خالية من الفقرات المتحيزة .

وبعد اطلاع الباحثة على الدراسات السابقة لم تجد دراسات أجريت على الكشف عن الأداء التفاضلي لمتغير الجنس في اختبار PISA في مجالي القراءة والعلوم لدورة 2012 باستخدام طريقتي مانتل هانزل وطريقة نسبة الأرجحية العظمى، ولذلك جاءت هذه الدراسة للكشف عن الأداء التفاضلي لمتغير الجنس في اختبار PISA للعام 2012.

الفصل الثالث

المنهجية والتصميم

تناول هذا الفصل وصفاً لمجتمع الدراسة وعينتها والإجراءات المتبعة.

1.3 مجتمع الدراسة:

تكون مجتمع الدراسة من جميع الطلبة الذين تقدموا لاختبار PISA 2012 في جميع الدول المشاركة والبالغ عددها(65) دولة.

2.3 عينة الدراسة:

تكونت عينة الدراسة من جميع الطلبة الذين تقدموا لإختبار PISA 2012 في المملكة الأردنية الهاشمية، والبالغ عددهم (7038) طالباً وطالبة من ضمن (233) مدرسة اختيرت عشوائياً من مدارس المملكة، وجميعهم من طلبة الصف العاشر، ويبين الجدول (2) توزيع عينة الدراسة حسب جنس المدرسة .

جدول (2)

توزيع عينة الدراسة حسب جنس المدرسة

جنس المدرسة	عدد الطلبة	النسبة المئوية للطلبة %	عدد المدارس	النسبة المئوية للمدرسة%
ذكور	3284	46.7	110	47.2
إناث	3530	50.2	115	49.4
مختلط	224	3.2	8	3.4
المجموع	7038	%100	233	%100

المركز الوطني لتنمية الموارد البشرية (التقرير الوطني PISA 2012)

3.3 أداة الدراسة

لتحقيق هدف الدراسة والمتمثل في الكشف عن الأداء التفاضلي لمتغير الجنس في اختبار PISA 2012 استخدمت نتائج طلبة الأردن الذين تقدموا للامتحان، وقد تم

اختيار جميع الإستجابات التي تحتوي على فقرات العلوم والبالغ عددها (53) فقرة، كذلك تم اختيار جميع الكراسات التي تحتوي على فقرات القراءة والبالغ عددها (44) فقرة، وفقرات الاختبارين مقسمة حسب الإجابة إلى اختيار من متعدد وإجابة مصاغه، يجيب عليها الطالب في زمن مقداره ساعتين، والجدولين(3) و(4) يبينان توزيع الفقرات للاختبارين على المستوى المعرفي .

الجدول(3)

المستوى المعرفي لكل فقرة من فقرات اختبار القراءة

رقم الفقرة	المستوى المعرفي للقراءة	رقم الفقرة	المستوى المعرفي للقراءة	رقم الفقرة	المستوى المعرفي للقراءة	رقم الفقرة	المستوى المعرفي للقراءة
1	وصول واسترجاع	13	ربط وتفسير	25	ربط وتفسير	37	ربط وتفسير
2	ربط وتفسير	14	ربط وتفسير	26	ربط وتفسير	36	وصول واسترجاع
3	ربط وتفسير	15	ربط وتفسير	27	ربط وتفسير	37	ربط وتفسير
4	ربط وتفسير	16	وصول واسترجاع	28	ربط وتفسير	38	ربط وتفسير
5	ربط وتفسير	17	تأمل وتقيم	29	وصول واسترجاع	39	وصول واسترجاع
6	ربط وتفسير	18	وصول واسترجاع	30	تأمل وتقيم	40	ربط وتفسير
7	تأمل وتقيم	19	ربط وتفسير	31	ربط وتفسير	41	ربط وتفسير
8	تأمل وتقيم	20	ربط وتفسير	32	تأمل وتقيم	42	وصول واسترجاع
9	ربط وتفسير	21	تأمل وتقيم	33	وصول واسترجاع	43	ربط وتفسير
10	ربط وتفسير	22	تأمل وتقيم	34	تأمل وتقيم	44	وصول واسترجاع
11	ربط وتفسير	23	ربط وتفسير	35	تأمل وتقيم		
12	وصول واسترجاع	24	تأمل وتقيم	36	وصول واسترجاع		

يتبين من الجدول أن (10) فقرات كانت لمستوى الوصول والاسترجاع، و(24) فقرة كانت لمستوى الربط والتفسير، و(10) فقرات لمستوى التأمل والتقييم.

جدول(4)

المستوى المعرفي لكل فقرة من فقرات اختبار العلوم

رقم الفقرة	المستوى المعرفي لفقرات العلوم	رقم الفقرة	المستوى المعرفي لفقرات العلوم	رقم الفقرة	المستوى المعرفي لفقرات العلوم	رقم الفقرة	المستوى المعرفي لفقرات العلوم
1	استخدام الأدلة العلمية	15	استخدام الأدلة العلمية	29	تحديد القضايا العلمية	43	استخدام الأدلة العلمية
2	تحديد القضايا العلمية	16	استخدام الأدلة العلمية	30	تحديد القضايا العلمية	44	شرح الظواهر علميا
3	شرح الظواهر علميا	17	شرح الظواهر علميا	31	استخدام الأدلة العلمية	45	استخدام الأدلة العلمية
4	شرح الظواهر علميا	18	شرح الظواهر علميا	32	شرح الظواهر علميا	46	استخدام الأدلة العلمية
5	شرح الظواهر علميا	19	تحديد القضايا العلمية	33	شرح الظواهر علميا	47	شرح الظواهر علميا
6	شرح الظواهر علميا	20	تحديد القضايا العلمية	34	تحديد القضايا العلمية	48	تحديد القضايا العلمية
7	استخدام الأدلة العلمية	21	استخدام الأدلة العلمية	35	استخدام الأدلة العلمية	49	شرح الظواهر علميا
8	استخدام الأدلة العلمية	22	شرح الظواهر علميا	36	تحديد القضايا العلمية	50	شرح الظواهر علميا
9	استخدام الأدلة العلمية	23	استخدام الأدلة العلمية	37	شرح الظواهر علميا	51	استخدام الأدلة العلمية
10	شرح الظواهر علميا	24	تحديد القضايا العلمية	38	استخدام الأدلة العلمية	52	شرح الظواهر علميا
11	شرح الظواهر علميا	25	استخدام الأدلة العلمية	39	شرح الظواهر علميا	53	شرح الظواهر علميا
12	شرح الظواهر علميا	26	استخدام الأدلة العلمية	40	تحديد القضايا العلمية		
13	شرح الظواهر علميا	27	شرح الظواهر علميا	41	تحديد القضايا العلمية		
14	تحديد القضايا العلمية	28	تحديد القضايا العلمية	42	استخدام الأدلة العلمية		

يتبين من الجدول أن (18) فقرة كانت لمستوى استخدام الأدلة العلمية، و(13) فقرة لمستوى تحديد القضايا العلمية، و(22) فقرة لمستوى شرح الظواهر علمياً.

4.3 صدق الاختبار

تم التحقق من صدق الاختبار من خلال إجراءات مضبوطة تتمثل بتحكيم الاختبار من قبل مختصين لغويين، ومختصين رياضيين، ومختصين في القياس والتقويم، لإخراج الاختبار بصورته الصحيحة، ويعتبر هذا الإجراء مؤشراً على صدق الاختبار.

وقد أجري التحليل العاملي على عينات من الدول المشاركة قبل البدء بتنفيذ الاختبارات الرسمية (OECD (PISA 2012.

5.3 ثبات الاختبار:

يتم تطبيق الاختبار بشكل مبدئي على عينة دولية عشوائية، ويتم التحقق من ثباته، وثبات فقراته والخصائص السيكومترية لفقراته ويعتبر هذا الإجراء مؤشراً على ثبات الاختبار.

6.3 إجراءات الدراسة

1. تم الحصول على كتاب تسهيل مهمة من جامعة مؤتة، وبناءاً عليه تم مخاطبة المركز الوطني لتنمية الموارد البشرية
2. تم الحصول على البيانات والتقارير النهائية لاختبار PISA 2012 من المركز الوطني لتنمية الموارد البشرية .
3. تحليل البيانات.

7.3 المعالجات الإحصائية

1. تم استخدام إحصائي مانتل -هانزل كاي تربيع (χ^2_{MH})، و إحصائي نسبة الأرجحية العظمى (χ^2_{LH}) للكشف عن الأداء التفاضلي لمتغير الجنس في فقرات اختبار PISA 2012 الدولي في العلوم والقراءة.
2. تم استخدام إحصائي(ت) للعينات المستقلة للكشف عن دلالة الفروق بين متوسطات أداء كل من الذكور والإناث
3. تم استخدام النسبة المئوية لحساب الأداء التفاضلي على أبعاد مستوى كل من الاختبارين

الفصل الرابع

عرض النتائج ومناقشتها والتوصيات

1.4 عرض النتائج ومناقشتها

يتناول هذا الفصل عرضاً للنتائج التي تم التوصل إليها وفقاً لتسلسل أسئلتها:

السؤال الأول: هل يوجد أداء تفاضلي لمتغير الجنس على فقـرات اختبـارPISA 2012 لمادة القراءة ؟

للإجابة عن السؤال تم حساب متوسط استجابات الإفراد على كل فقـرة مـن فقرات اختبار القراءة، كما يبينها الجدول (5) .

الجدول(5)

متوسطات استجابات الأفراد على كل فقرة من فقرات اختبار القراءة

الفقرة	الإناث: الوسط	الإناث: الانحراف المعياري	الذكور: الوسط	الذكور: الانحراف المعياري	المجموع: الوسط	المجموع: الانحراف المعياري	الفقرة	الإناث: الوسط	الإناث: الانحراف المعياري	الذكور: الوسط	الذكور: الانحراف المعياري	المجموع: الوسط	المجموع: الانحراف المعياري
1	0.100	0.300	0.057	0.233	0.079	0.271	23	0.711	0.453	0.532	0.499	0.624	0.484
2	0.417	0.493	0.326	0.469	0.373	0.484	24	0.545	0.498	0.345	0.476	0.448	0.497
3	0.422	0.494	0.293	0.455	0.359	0.480	25	0.020	0.141	0.014	0.119	0.017	0.130
4	0.631	0.483	0.418	0.494	0.528	0.499	26	0.431	0.495	0.294	0.456	0.365	0.482
5	0.344	0.475	0.257	0.437	0.301	0.459	27	0.391	0.488	0.334	0.472	0.363	0.481
6	0.104	0.306	0.089	0.284	0.097	0.296	28	0.214	0.410	0.112	0.316	0.165	0.371
7	0.244	0.430	0.095	0.293	0.171	0.377	29	0.877	0.329	0.675	0.469	0.778	0.415
8	0.216	0.412	0.096	0.295	0.158	0.365	30	0.535	0.499	0.335	0.472	0.438	0.496
9	0.709	0.454	0.451	0.498	0.583	0.493	31	0.620	0.485	0.463	0.499	0.544	0.498
10	0.258	0.438	0.175	0.380	0.218	0.413	32	0.447	0.497	0.270	0.444	0.362	0.481
11	0.572	0.495	0.349	0.477	0.463	0.499	33	0.356	0.479	0.166	0.372	0.264	0.441
12	0.764	0.425	0.575	0.495	0.673	0.469	34	0.671	0.470	0.381	0.486	0.531	0.499
13	0.561	0.496	0.411	0.492	0.489	0.500	35	0.222	0.416	0.107	0.309	0.166	0.372
14	0.419	0.494	0.446	0.497	0.432	0.495	36	0.844	0.363	0.645	0.479	0.747	0.435
15	0.254	0.435	0.092	0.289	0.176	0.381	37	0.325	0.468	0.228	0.420	0.278	0.448
16	0.522	0.500	0.339	0.474	0.434	0.496	38	0.041	0.199	0.030	0.170	0.036	0.185
17	0.373	0.484	0.222	0.416	0.300	0.459	39	0.946	0.227	0.793	0.405	0.872	0.334
18	0.710	0.454	0.484	0.500	0.601	0.490	40	0.706	0.456	0.551	0.498	0.631	0.483
19	0.527	0.499	0.278	0.448	0.407	0.491	41	0.787	0.410	0.552	0.498	0.673	0.469
20	0.388	0.488	0.309	0.462	0.350	0.477	42	0.198	0.399	0.101	0.302	0.151	0.358
21	0.607	0.489	0.482	0.500	0.546	0.498	43	0.100	0.300	0.091	0.287	0.095	0.294
22	0.675	0.468	0.517	0.500	0.599	0.490	44	0.633	0.482	0.423	0.494	0.531	0.499

يلاحظ وجود فروق ظاهرية بين متوسطات استجابات كل من الذكور والإناث على فقرات الاختبار، وللكشف عن أفضلية الأداء على كل من فقرات الاختبار تـم استخدام كل من اختبار مانتل هانزل واختبـار الأرجحيـة واختبـار(ت) للعينـات المستقلة، ويبين الجدول (6) نتائج ذلك.

الجدول (6)

اختبار نسبة الأرجحية، مانتل هانزل، واختبار(ت) للكشف عن الأداء التفاضلي في اختبار القراءة.

الفقرة	نسبة الأرجحية χ^2_{LH}	الدلالة	مانتـل هانـزل χ^2_{MH}	الدلالة	اختبـار T قيمة(t)	الدلالة	الأداء التفاضلي لصالح
1	12.681	0.000	11.889	0.001	3.541	0.000	الإناث
2	17.715	0.000	17.268	0.000	4.219	0.000	الإناث
3	36.069	0.000	35.290	0.000	6.041	0.000	الإناث
4	97.966	0.000	96.347	0.000	10.089	0.000	الإناث
5	19.342	0.000	18.845	0.000	4.407	0.000	الإناث
6	1.499	0.221	1.321	0.250	1.223	0.221	لا يوجد
7	86.409	0.000	82.390	0.000	9.318	0.000	الاناث
8	58.792	0.000	56.293	0.000	7.666	0.000	الإناث
9	145.809	0.000	143.012	0.000	12.433	0.000	الإناث
10	21.333	0.000	20.686	0.000	4.624	0.000	الإناث
11	105.114	0.000	103.210	0.000	10.464	0.000	الإناث
12	84.124	0.000	82.650	0.000	9.326	0.000	الإناث
13	46.223	0.000	45.422	0.000	6.859	0.000	الإناث
14	1.528	0.216	1.419	0.234	-1.236	0.217	لايوجد
15	96.711	0.000	91.660	0.000	9.855	0.000	الإناث
16	72.591	0.000	71.282	0.000	8.633	0.000	الاناث
17	57.422	0.000	56.018	0.000	7.634	0.000	الإناث
18	112.922	0.000	110.991	0.000	10.872	0.000	الإناث
19	136.850	0.000	133.716	0.000	11.996	0.000	الإناث
20	14.658	0.000	14.270	0.000	3.835	0.000	الإناث
21	33.663	0.000	33.068	0.000	5.839	0.000	الإناث
22	55.801	0.000	54.901	0.000	7.551	0.000	الإناث
23	69.705	0.000	68.539	0.000	8.468	0.000	الإناث
24	82.505	0.000	81.002	0.000	9.231	0.000	الإناث
25	1.049	0.306	0.721	0.396	1.020	0.308	لايوجد
26	41.089	0.000	40.226	0.000	6.452	0.000	الإناث
27	6.807	0.009	6.553	0.010	2.610	0.009	الإناث
28	38.391	0.000	36.893	0.000	6.191	0.000	الإناث

الفقرة	نسبة الأرجحية χ^2_{LH}	الدلالة	مانتـل هانـــزل χ^2_{MH}	الدلالة	اختبـــارT قيمة(t)	الدلالة	الأداء التفاضلي لصالح
29	123.704	0.000	119.680	0.000	11.330	0.000	الإناث
30	83.347	0.000	81.805	0.000	9.277	0.000	الإناث
31	52.540	0.000	51.689	0.000	7.324	0.000	الإناث
32	71.847	0.000	70.302	0.000	8.576	0.000	الإناث
33	98.221	0.000	94.971	0.000	10.027	0.000	الإناث
34	177.264	0.000	173.591	0.000	13.809	0.000	الإناث
35	51.205	0.000	49.231	0.000	7.160	0.000	الإناث
36	111.704	0.000	108.756	0.000	10.764	0.000	الإناث
37	24.283	0.000	23.654	0.000	4.940	0.000	الإناث
38	2.072	0.150	1.731	0.188	1.434	0.152	لايوجد
39	114.148	0.000	107.565	0.000	10.715	0.000	الإناث
40	54.686	0.000	53.783	0.000	7.475	0.000	الإناث
41	132.642	0.000	129.828	0.000	11.813	0.000	الإناث
42	39.019	0.000	37.466	0.000	6.238	0.000	الإناث
43	0.517	0.472	0.414	0.520	0.718	0.473	لايوجد
44	91.958	0.000	90.419	0.000	9.768	0.000	الإناث

يتضح من الجدول(6) وجود أداء تفاضلي لصالح الإناث على (39) فقرة من فقرات الاختبار والتي كان عددها (44) وباتفاق جميع الاختبارات المستخدمة، كمـا أن(5) فقرات لم تظهر أفضلية في الأداء لأي من الجنسين .

ويشر ذلك إلى وجود فجوة جندرية في التعليم، ففي السنوات الأخيرة اتـسعت الفجوة بين الجنسين في اختبارات القراءة، وذلك لتغير اتجاهات الطلبة وسـلوكياتهم (تقرير منظمة OECD 2009).

وقد يعزى ذلك إلى الجدية التي تبديها الإناث عند الاستجابة على الاختبارات، وميل الطالبات للقراءة والمطالعة على خلاف الذكور الذين يبـدون مـيلا للترفيـه ومشاركة الأقران في الرحلات والمناسبات في هذه المرحلة العمرية، وهـذا يتفـق أيضا مع العديد من الدراسات التي أظهرت تفوق الإناث على الذكور فـي القـدرة اللفظية والقراءة في هذه المرحلة العمرية، منها دراسة (البستنجي، 2004) .

السؤال الثاني:" هل يوجـد أداء تفاضـلي لمتغيـر الجـنس علـى فقـرات اختبارPISA 2012 لمادة العلوم ؟"

للإجابة عن هذا السؤال تم حساب متوسط استجابات الإفراد على كل فقرة من فقرات اختبار العلوم كما يبينها الجدول (7) .

الجدول(7)

متوسطات استجابات الأفراد على كل فقرة من فقرات اختبار العلوم

الفقرة	الإناث		الذكور		المجموع		الفقرة	الإناث		الذكور		المجموع	
	الوسط	الانحراف المعياري	الوسط	الانحراف المعياري	الوسط	الانحراف المعياري		الوسط	الانحراف المعياري	الوسط	الانحراف المعياري	الوسط	الانحراف المعياري
1	0.402	0.491	0.214	0.411	0.311	0.463	28	0.639	0.480	0.472	0.499	0.558	0.497
2	0.165	0.372	0.089	0.285	0.128	0.335	29	0.531	0.499	0.385	0.487	0.460	0.499
3	0.880	0.325	0.744	0.437	0.814	0.389	30	0.241	0.428	0.132	0.338	0.188	0.391
4	0.455	0.498	0.341	0.474	0.400	0.490	31	0.364	0.481	0.287	0.452	0.326	0.469
5	0.371	0.483	0.247	0.432	0.311	0.463	32	0.459	0.499	0.393	0.489	0.427	0.495
6	0.148	0.355	0.170	0.376	0.159	0.365	33	0.222	0.416	0.249	0.433	0.235	0.424
7	0.369	0.483	0.227	0.419	0.300	0.458	34	0.517	0.500	0.434	0.496	0.477	0.500
8	0.455	0.498	0.249	0.433	0.355	0.479	35	0.291	0.455	0.349	0.477	0.319	0.466
9	0.373	0.484	0.320	0.467	0.348	0.476	36	0.443	0.497	0.278	0.448	0.363	0.481
10	0.102	0.303	0.090	0.286	0.096	0.295	37	0.339	0.474	0.255	0.436	0.299	0.458
11	0.387	0.487	0.261	0.439	0.326	0.469	38	0.375	0.484	0.288	0.453	0.333	0.471
12	0.197	0.398	0.169	0.375	0.183	0.387	39	0.525	0.500	0.396	0.489	0.462	0.499
13	0.348	0.477	0.302	0.459	0.326	0.469	40	0.331	0.471	0.275	0.447	0.303	0.460
14	0.308	0.462	0.264	0.441	0.287	0.452	41	0.368	0.483	0.259	0.438	0.315	0.465
15	0.226	0.418	0.223	0.417	0.224	0.417	42	0.634	0.482	0.390	0.488	0.516	0.500
16	0.669	0.471	0.491	0.500	0.583	0.493	43	0.640	0.480	0.466	0.499	0.555	0.497
17	0.264	0.441	0.196	0.397	0.231	0.422	44	0.293	0.455	0.299	0.458	0.296	0.456
18	0.687	0.464	0.478	0.500	0.585	0.493	45	0.402	0.491	0.217	0.413	0.312	0.463
19	0.483	0.500	0.376	0.485	0.431	0.495	46	0.261	0.440	0.156	0.363	0.210	0.408
20	0.288	0.453	0.219	0.414	0.254	0.435	47	0.417	0.493	0.420	0.494	0.418	0.493
21	0.310	0.463	0.335	0.472	0.322	0.467	48	0.087	0.283	0.058	0.235	0.073	0.261
22	0.490	0.500	0.378	0.485	0.436	0.496	49	0.311	0.463	0.308	0.462	0.310	0.462
23	0.203	0.402	0.107	0.309	0.156	0.363	50	0.849	0.358	0.711	0.454	0.782	0.413
24	0.481	0.500	0.412	0.493	0.448	0.497	51	0.097	0.296	0.078	0.269	0.088	0.283
25	0.273	0.446	0.281	0.450	0.277	0.448	52	0.520	0.500	0.445	0.497	0.484	0.500
26	0.434	0.496	0.332	0.471	0.384	0.487	53	0.366	0.482	0.371	0.483	0.368	0.482
27	0.302	0.459	0.203	0.402	0.254	0.435							

يلاحظ من الجدول(7) وجود فروق ظاهرية بين متوسطات استجابات كل من الذكور والإناث على فقرات الاختبار وللكشف عن أفضلية الأداء على كل من فقرات

الاختبار تم استخدام كل من اختبار مانتل هانزل واختبار الأرجحيـة واختبـار (ت) للعينات المستقلة ونتائج كل منها في الجدول (8) تبين ذلك .

الجدول (8)

اختبار نسبة الأرجحية، مانتل هانزل، واختبار(ت) للكشف عن الأداء التفاضلي في اختبار العلوم.

الفقرة	نسبة الارجحية		مانتل هانزل		اختبارT		الأداء التفاضلي
	χ^2_{LH}	الدلالة	χ^2_{MH}	الدلالة	قيمة(t)	الدلالة	لصالح
1	89.13	0.000	86.844	0.000	9.562	0.000	الإناث
2	28.211	0.000	26.994	0.000	5.293	0.000	الإناث
3	65.646	0.000	63.883	0.000	8.172	0.000	الإناث
4	28.684	0.000	28.093	0.000	5.379	0.000	الإناث
5	38.137	0.000	37.261	0.000	6.205	0.000	الإناث
6	2.023	0.155	1.857	0.173	-1.423	0.155	لايوجد
7	51.553	0.000	50.347	0.000	7.228	0.000	الإناث
8	99.96	0.000	97.663	0.000	10.163	0.000	الإناث
9	6.661	0.010	6.415	0.011	2.582	0.010	الإناث
10	0.854	0.355	0.722	0.396	0.923	0.356	لايوجد
11	38.516	0.000	37.661	0.000	6.238	0.000	الإناث
12	2.761	0.097	2.571	0.109	1.660	0.097	لايوجد
13	5.148	0.023	4.93	0.026	2.269	0.023	الإناث
14	5.017	0.025	4.793	0.029	2.240	0.025	الإناث
15	0.016	0.900	0.005	0.941	0.126	0.900	لايوجد
16	68.718	0.000	67.62	0.000	8.403	0.000	الإناث
17	13.594	0.000	13.143	0.000	3.688	0.000	الإناث
18	93.118	0.000	91.536	0.000	9.833	0.000	الإناث
19	24.407	0.000	23.895	0.000	4.961	0.000	الإناث
20	12.985	0.000	12.568	0.000	3.606	0.000	الإناث
21	1.391	0.238	1.282	0.258	-1.180	0.238	لايوجد
22	26.722	0.000	26.179	0.000	5.193	0.000	الإناث
23	36.077	0.000	34.637	0.000	5.997	0.000	الإناث
24	9.808	0.002	9.517	0.002	3.136	0.002	الإناث
25	0.178	0.673	0.14	0.709	-0.422	0.673	لايوجد
26	23.628	0.000	23.108	0.000	4.878	0.000	الإناث
27	27.569	0.000	26.831	0.000	5.263	0.000	الإناث
28	59.565	0.000	58.602	0.000	7.809	0.000	الإناث
29	45.096	0.000	44.303	0.000	6.772	0.000	الإناث
30	40.897	0.000	39.53	0.000	6.405	0.000	الإناث
31	14.737	0.000	14.336	0.000	3.858	0.000	الإناث

الفقرة	نسبة الارجحية		مانتل هانزل		اختبارT		الأداء التفاضلي
	χ^2_{LH}	الدلالة	χ^2_{MH}	الدلالة	قيمة(t)	الدلالة	لصالح
32	9.565	0.002	9.282	0.002	3.096	0.002	الإناث
33	2.224	0.136	2.074	0.150	-1.492	0.136	لايوجد
34	14.355	0.000	13.997	0.000	3.798	0.000	الإناث
35	7.995	0.005	7.726	0.005	-2.832	0.005	الذكور
36	60.737	0.000	59.475	0.000	7.872	0.000	الإناث
37	17.875	0.000	17.394	0.000	4.235	0.000	الإناث
38	18.012	0.000	17.553	0.000	4.253	0.000	الإناث
39	35.538	0.000	34.887	0.000	5.999	0.000	الإناث
40	7.733	0.005	7.451	0.006	2.782	0.005	الإناث
41	28.942	0.000	28.265	0.000	5.399	0.000	الإناث
42	125.919	0.000	123.645	0.000	12.449	0.000	الإناث
43	64.396	0.000	63.369	0.000	8.127	0.000	الإناث
44	0.107	0.743	0.078	0.780	-0.328	0.743	لايوجد
45	84.417	0.000	82.329	0.000	9.305	0.000	الإناث
46	35.213	0.000	34.131	0.000	4.875	0.000	الإناث
47	0.02	0.888	0.009	0.923	-0.141	0.888	لايوجد
48	6.485	0.011	5.999	0.014	2.537	0.011	الإناث
49	0.012	0.913	0.004	0.950	0.110	0.913	لايوجد
50	59.116	0.000	57.774	0.000	7.761	0.000	الإناث
51	2.258	0.133	2.023	0.155	1.500	0.134	لايوجد
52	11.639	0.001	11.324	0.001	3.418	0.001	الإناث
53	0.045	0.832	0.028	0.867	-0.213	0.832	لايوجد

يتضح من الجدول(8) وجود أداء تفاضلي لصالح الإناث على (40) فقرة من فقرات الاختبار، وفقرة واحده لصالح الذكور، و (12) فقرة لم تظهر أفضلية في الأداء لأي من الجنسين .

ويمكن تفسير ذلك من خلال الجدية التي تبديها الإناث عند الاستجابة على الاختبارات الدولية، ولأن ميل الذكور للفروع المهنية أكثر من الفروع الأكاديمية في هذه المرحلة .

والفقرة (35) التي أظهرت أداء تفاضليا للذكور هي (يعد الأكسجين أحد الغازات التي تخرج من عادم السيارات.

أي العبارات التالية أدناه توضح السبب ؟

أ- ينتج الأكسجين من محرك السيارة.

ب - ينتج الأكسجين من التفاعلات في المحول الحفري.

ج - لا يستخدم كل الأكسجين الذي يذهب إلى حجرة المحرك.

د – يعطينا الوقود الحديث أكسجين أكثر من الأكسجين المستخدم.

ويمكن تفسير ذلك من خلال طبيعة محتوى الفقرة التي تناولت موضوع يخص السيارات، والذي يجذب انتباه الذكور خاصة في هذه المرحلة العمرية .

السؤال الثالث: ما هي أبعاد المستوى المعرفي التي تظهر أداء تفاضليا على فقرات اختبار القراءة في اختبار PISA 2012؟

للإجابة على السؤال تم حساب نسبة الفقرات ذات الأداء التفاضلي في كل من أبعاد اختبار القراءة والمكون من ثلاثة أبعاد (وصول واسترجاع، ربط وتفسير، تأمل وتقييم) والجدول (9) يبين ذلك.

الجدول(9)

نسبة الأداء التفاضلي في أبعاد المستوى المعرفي لاختبار القراءة .

المستوى المعرفي	عدد الفقرات	أداء تفاضلي على البعد (عدد الفقرات)	نسبة الأداء التفاضلي على البعد
وصول واسترجاع	10	10	100%
ربط وتفسير	24	19	79%
تأمل وتقييم	10	10	100%

يتبين من الجدول (9) أن الأداء التفاضلي كان في الربط والتفسير بنسبة 79%، أما في الوصول والاسترجاع والبعد الثالث التأمل والتقييم كان بنسبة 100% وهذا يشير إلى أن المستويات المعرفية لمبحث العلوم في الأردن غير متوافقة مـع المستويات المعرفية المناظرة في اختبار PISA فهي بحاجة إلى مزيد من التطـوير والتعديل لتصبح أكثر انسجاماً واتصالاً مع اختبار PISA

السؤال الرابع: ماهي أبعاد المستوى المعرفي التي تظهر أداء تفاضليا على فقرات اختبار العلوم في اختبار PISA 2012 ؟

للإجابة على السؤال تم حساب نسبة الفقرات ذات الأداء التفاضلي في كل من أبعاد المستوى المعرفي لاختبار العلوم والمكون من ثلاثة أبعاد (استخدام الأدلة العلمية، تحديد القضايا العلمية، وشرح الظواهر علميا)، والجدول (10) يبين ذلك.

الجدول(10)

نسبة الأداء التفاضلي في أبعاد المستوى المعرفي لاختبار العلوم:

المستوى المعرفي	عدد الفقرات	أداء تفاضلي على البعد (عدد الفقرات)	نسبة الأداء التفاضلي على البعد
استخدام الأدلة العلمية	18	14	%78
تحديد القضايا العلمية	13	13	%100
شرح الظواهر علميا	22	14	%64

يتبين من الجدول (10) أن نسبة الأداء التفاضلي في بعد استخدام الأدلة العلمية 78%، وفي بعد شرح الظواهر علميا 64%، وفي بعد تحديد القضايا العلمية 100% .

أي أن الطلبة لديهم مشكلة وقدرة منخفضة في تحديد القضايا العلمية واستخدام الأدلة العلمية، وهذا يشير إلى ضرورة تطوير وإثراء مناهج العلوم بحيث تشتمل وتركز على هذه المستويات المعرفية في هذه المرحلة ، وتوفير بيئات مدرسية مناسبة لحث الطلبة على تحديد القضايا العلمية، واستخدام الأدلة العلمية، وشرح الظواهر علمياً.

2.4 التوصيات

توصي الدراسة بما يلي:

1) الاستفادة من نتائج دراسة البرنامج الدولي (PISA) ونتائج الدراسات التقييمية الأخرى لتطوير مشروع إصلاح تربوي وطني شامل يركز على القضايا التربوية الرئيسية ويهتم بالجوانب النوعية.
2) زيادة تطلع وزارة التربية والتعليم في الأردن لردم الفجوة الجندرية في التعليم.
3) تغيير البرامج والممارسات التربوية والتعليمية لتحقيق مخرجات تعليمية تتماشى مع الاختبارات الوطنية والدولية .

4) إثراء مناهج العلوم في المرحلة الأساسية في الأردن بمستويات تحاكي مستويات اختبار PISA.

5) إجراء المزيد من الدراسات في الكشف عن الأداء التفاضلي لمتغيرات أخـرى في اختبار PISA.

المراجع

أ) المراجع العربية

البرصان، إسماعيل سلامه.(2013).الأداء التفاضلي لمتغير الجنس لفقرات الاختبار الوطني الأردني لضبط نوعية التعلم لمادة الرياضيات للصف العاشـر، **مجلة كلية التربية** بالزقازيق،28(79)1-35.

البستنجي، محمود. (2004).**مقارنة أربع طرق للكشف عن الأداء التفاضلي لمتغير الجنس في فقرات قدرات عقلية خاصة للفئة العمرية (15-16) سـنة في الأردن،** رسالة دكتوراه غير منشورة، جامعة عمان العربية، عمـان، الأردن.

جرادات، عبدا . (2003)**. مقارنة بين طريقتي مانتل هانزل وطريقـة الـصعوبة المحولة في الكشف عن تحيز الفقرات،** رسالة ماجستير غير منـشورة، جامعة مؤتة، الكرك، الأردن.

حمادنه، إياد.(2007). **الأداء التفاضلي لفقرات اختبارات"تحديد الكفاءة اللغوية في اللغة الإنجليزية في الجامعات الأردنية": دراسـة مقارنـة،** أطروحـة دكتوراة غير منشورة،جامعة اليرموك، اربد، الأردن.

الطراونة، محمود زعل.(2011). **تحري التحيز في فقرات الاختبار الوطني لطلبة الصف الثامن في مادتي الرياضيات والعلوم (2010)**، رسالة ماجستير غير منشورة، جامعة مؤتة،الأردن.

العطيوي، إيمان. (2004). **الكشف عن الأداء التفاضلي لمتغير الجـنس لفقـرات اختبار تحصيلي في العلوم العامة للصف الثامن الأساسي فـي مديريـة تربية عمان الرابعة،** رسالة ماجستير غيـر منـشورة، جامعـة عمـان العربية، عمان، الأردن.

علام، صلاح الدين.(2006).**القياس والتقـويم التربـوي والنفـسي،** دار الفكـر العربي، القاهرة، مصر .

عوده، أحمد. (2004). **القياس والتقويم في العملية التدريسية.** اربد: دار الأمـل للنشر.

المركز الوطني لتنمية الموارد البشرية.(2012).**التقرير الوطني الأردني عن الدراسة الدولية بيزا لعام 2012**. وحدة المتابعة والتخطيط، عمان.

المساعفة، أحمد جميل.(2005). **درجة تمثيل كتب الرياضيات المدرسية للصفوف (من الرابع وحتى الثامن) في الأردن للمفاهيم الأساسية ولشكل ومستويات الأسئلة في اختبار (TIMSS-R)** ، رسالة ماجستير غير منشورة، الجامعة الأردنية، عمان ، الأردن.

النبهان، موسى. (2004). **أساسيات القياس في العلوم السلوكية**، دار الشروق للنشر والتوزيع: عمان، الأردن.

الهريشات، محمد. (2004). **تحري التحيز في فقرات ايسر للتفكير الاستدلالي**، رسالة ماجستير غير منشورة، جامعة مؤتة، الكرك، الأردن.

ب) المراجع الأجنبية

Camilli, G., and Shepard, L. (1994). **Methods for identifying biased test items**. Thousand Oaks, CA: Sage Publications.

Crocker, L and Algina,J.(1986).**Introduction to classical and modern theory**.Holt Rinehart and Winston.

Dorans, Neil J & Holland, Paul W. (1993). **DIF Detection and Description Mantel – Haenszel and Standardization**. N. J Educational Testing Service Princeton.

Hambleton ,R & Rogers, J.(1995) **Item bias review Practical Assessment, Research and Evaluation** . Retrieved May 15, 2009.

Hambleton , R and Swaminithan , H. & Rogers,J (1991) . **Fundamentals of item response theory** . Sage publications , Newbury perk .

Hamilton,L and Snow,T & Cresst,H.(1998).Exploring Differential item functioning on science achievement – Tests**, Applied measurement in education**,12,(3),191-212

Hamilton, L. (1999). Detecing Gender- Based Differential item Functioning On constructed response science test. **Applied measurement in education**, 12, (3), 30-56.

Hidalgo, M & Lopez, P. (2002). Twostage equating in differential item functioning detection under the graded response model with the Raju area measures and the lord statistic. **Educational and Psychological Measurement**, 62(1), 32–44.

Hoover, H. & Kolen, M.(1984).The Reliability of sex item bias indices. **Applied Psychological Measurement**, 8(2). 172-180.

Ironson ,G(1979). **A Comparative study of several Methods of assessing item bias**.(Unpublished Doctoral Dissertation, University of Wisconsin Madison)

Jensen, A (1980). **Bias in mental Testing**. NewYork,A Deivision of Macmillan Publishing Co.Inc.

Katherine,R & Shuwan,C. (2001).An Examination of item context effects,DIF, and gender DIF,**Applied Measurement in Education** 14(1),73-90

Kim,S & Cohen,A.S .(1995). A comparison of lord'S Chi square Raju,S Area Measures ,And likelihood Ratio Test On Detection of Differential Item Function ,**Applied Measurements In Education** ,4(3),291-312

Le, Luc T.(2009). Investigating Gender Differential Item Functioning across Countries and Test Languages for PISA Science items, **International Journal of Testing**, 9 (2), 122 -133.

Nandakumer,R (1993)**.**A FORTRAN 77 Program for Detecting Differentional Item Function Through the Mantel -Haenszel Statistics, **Educational and Psychological Measurement**,53, (3),210-235.

Ning, W & Lane, S. (1996). Detection of gender – related differential item functioning in mathematics performance assessment. **Applied Measurement in Education**. Vol 9. 171-211**OECD(2012).**PISA 2012 Results**. http://www.oecd.org/pisa/keyfindings/pisa-2012-results.htm.**

OECD (2012) .**PISA 2012 Assessment Framework – Key competencies in reading , mathematics and science** .

Osterlind,J.(1983).**Test Item Bias**. Beverly Hills: sage publication.

Raju, N and Drassgow , F& Slind , J.(1993). An Empirical Comparison Of The Area Methods, Lords Chi-Squar Test, And The Mantel-Haenszel Technique For Assessing DIF**, Educational and Psychological Measurement**,53(1),301-314.

Scheuneman, J. (1979). A method of assessing bias in test item**. Journal of Eductional Measuremen**t. 16, (3).

Skaggas, G and Lissits, R .(1992) .A method of assessing bias in test item across different test administration :implications of another failure .**journal of educational measurements**,29,(3),227-242.

Swaminathan, H. Rogers, J. H.(1990). Detecting Differential Item Functioning Using Logistic Regression Procedures **, Journal of Educational Measurement,** 27, (4), pages 361–370

Thorndike, Robert. (1982).**Applied Psychometrics**. Boston Houghton Mifflin Company.

Williams, S, (1977).The "unbiased" anchor :bridging the gap between DIF and item bias. **Applied Measurements in Education**, 10(3),253-267

YAN, Song (2005). **Gender-related Differential Item Functioning in Mathematics Assessment on the Third International Mathematics and Science Study-Repeat (TIMSS-R)**. The University of Toledo, ProQuest Dissertations Publishing, 2005. 3177610.

Zumbo, D. B. (1999). **A handbook on the theory and methods of differential ıtem functioning (DIF): Logistic regression modeling as a unitary framework for binary and likert-type item scores**. Ottowa: Directorate of Human Resources Research and Evaluation, Department of National Defense.

الملحق(أ)

كتاب تسهيل المهمة

جامعة مؤتة

مكتب الرئيس



MU'TAH UNIVERSITY

President Office

الرقم: ٨٠٠/ ٥٧١٩

التاريخ: ١٣/محرم/١٤٣٧هـ

الموافق: ٢٦/١٠/٢٠١٥م

Ref.:

Date:

السادة مركز تطوير الموارد البشرية المحترمين

عمّان

تحية طيبة، وبعد؛

فأرجو التكرم بالموافقة والإيعاز لمن يلزم؛ لتسهيل مهمة الطالبة خالدة محمود الشبايشة/ كلية العلوم التربوية، في الحصول على المعلومات والبيانات اللازمة لإعداد دراستها الموسومة بـ: "**الكشف عن الأداء التفاضلي لمتغير الجنس في اختبار Pisa لعام ٢٠١٢**"، من المعنيين لديكم؛ وذلك استكمالاً لمتطلبات الحصول على درجة الماجستير.

شاكرين لكم اهتمامكم وحرصكم على التعاون مع جامعة مؤتة.

وتفضلوا بقبول فائق الاحترام،،،

رئيس الجامعة

نائب الرئيس لشؤون الكليات الإنسانية

وشؤون الطلبة والمجتمع

أ.د. عبد الحميد إبراهيم المجالي

نسخة/ عميد كلية الدراسات العليا

مؤتة - الكرك - الأردن - هاتف: ٢٣٧٢٣٨٠-٣-٩٦٢+ ص.ب: (٧) الرمز البريدي: (٦١٧١٠) فاكس: ٢٣٧٥٥٤٠-٣-٩٦٢+

Mu'tah-Karak-Jordan-Tel: +962-3-2372380 P.O.Box: (7) Zip Code: (61710) Fax: +962-3-2375540

www.mutah.edu.jo E-mail: mutah@mutah.edu.jo